DOCUMENT RESUME

ED 249 214 SP 025 342

AUTHOR Hofmeister, Alan M.; Preston, Caroline N.
TITLE Curriculum-Based Assessment and Evaluation Procedures. [Revised].
INSTITUTION Minnesota Univ., Minneapolis. National Support Systems Project.
SPONS AGENCY Office of Special Education and Rehabilitative Services (ED), Washington, DC. Div. of Personnel Preparation.
PUB DATE Sep 81
GRANT OEG007902045
NOTE 125p.; For other modules in this series, see ED 238 844 and SP 025 332-354. For the genesis of these modules, see ED 186 399. Reviewed by Maynard C. Reynolds and Thomas Oakland.
PUB TYPE Guides - Classroom Use - Guides (For Teachers) (052)

EDRS PRICE MF01/PC05 Plus Postage.
DESCRIPTORS *Curriculum Evaluation; Educational Diagnosis; *Evaluation Methods; Higher Education; Learning Modules; *Mainstreaming; Preservice Teacher Education; Student Evaluation; Student Needs; *Teacher Education Curriculum; *Teacher Educators
IDENTIFIERS Education for All Handicapped Children Act

ABSTRACT

This module (part of a series of 24 modules) is on educational assessment and its purposes. The genesis of these materials is in the 10 "clusters of capabilities," outlined in the paper, "A Common Body of Practice for Teachers: The Challenge of Public Law 94-142 to Teacher Education." These clusters form the proposed core of professional knowledge needed by teachers in the future. The module is to be used by teacher educators to reexamine and enhance their current practice in preparing classroom teachers to work competently and comfortably with children who have a wide range of individual needs. The module includes objectives, scales for assessing the degree to which the identified knowledge and practices are prevalent in an existing teacher education program, and self-assessment test items. Bibliographic references and journal articles are included on student assessment and educational diagnosis. (JD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED249214

CURRICULUM-BASED ASSESSMENT AND EVALUATION PROCEDURES

Prepared by

Alan M. Hofmeister

Caroline N. Preston

Reviewed by

Maynard C. Reynolds

Thomas Oakland

SEPTEMBER 1981

SP 025 342

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

Extending the Challenge:

Working Toward a Common Body of Practice for Teachers

Concerned educators have always wrestled with issues of excellence and professional development. It is argued in the paper, "A Common Body of Practice for Teachers: The Challenge of Public Law 94-142 to Teacher Education,"* that the Education for All Handicapped Children Act of 1975 provides the necessary impetus for a concerted reexamination of teacher education. Further, it is argued that this reexamination should enhance the process of establishing a body of knowledge common to the members of the teaching profession. The paper continues, then, by outlining clusters of capabilities that may be included in the common body of knowledge. These clusters of capabilities provide the basis for the following materials.

The materials are oriented toward assessment and development. First, the various components, rating scales, self-assessments, sets of objectives, and respective rationale and knowledge bases are designed to enable teacher educators to assess current practice relative to the knowledge, skills, and commitments outlined in the aforementioned paper. The assessment is conducted not necessarily to determine the worthiness of a program or practice, but rather to reexamine current practice in order to articulate essential common elements of teacher education. In effect then, the "challenge" paper and the ensuing materials incite further discussion regarding a common body of practice for teachers.

Second, and closely aligned to assessment, is the developmental perspective offered by these materials. The assessment process allows the user to view current practice on a developmental continuum. Therefore, desired or more appropriate practice is readily identifiable. On another, perhaps more important dimension, the "challenge" paper and these materials focus discussion on preservice teacher education. In making decisions regarding a common body of practice, it is essential that specific knowledge, skill and commitment be acquired at the preservice level. It is also essential that other additional specific knowledge, skill, and commitment be acquired as a teacher is inducted into the profession and matures with years of experience. Differentiating among these levels of professional development is paramount. These materials can be used in forums in which focused discussion will explicate better the necessary elements of preservice teacher education. This explication will then allow more productive discourse on the necessary capabilities of beginning teachers and the necessary capabilities of experienced teachers.

---

*Published by the American Association of Colleges for Teacher Education, Washington, DC, 1980 ($5.50)

In brief, this work is an effort to capitalize on the creative ferment of the teaching profession in striving toward excellence and professional development. The work is to be viewed as evolutionary and formative. Contributions from our colleagues are heartily welcomed.

This paper presents one module in a series of resource materials which are designed for use by teacher educators. The genesis of these materials is in the ten "clusters of capabilities," outlined in the paper, "A Common Body of Practice for Teachers: The Challenge of Public Law 94-142 To Teacher Education," which form the proposed core of professional knowledge needed by professional teachers who will practice in the world of tomorrow. The resource materials are to be used by teacher educators to reexamine and enhance their current practice in preparing classroom teachers to work competently and comfortably with children who have a wide range of individual needs. Each module provides further elaboration of a specified "cluster of capabilities"--in this case, assessment and evaluation procedures. This particular module is a revision of an earlier module developed by the authors for the School Psychology Inservice Training Network, Department of Psychoeducational Studies, University of Minnesota, directed by James Ysseldyke, and Richard Weinberg. This module was one of several originally developed under the editorship of Thomas Oakland, University of Texas, on the topic of non-biased assessment.

ASSESSMENT AND EVALUATION PROCEDURES

The competency domain which includes knowledge and skill in "assessment and evaluation procedures" is one of critical importance in fulfilling both the mandate and spirit of Public Law 94-142. In the past, teachers have tended to function as recipients of assessment and evaluation results. These results were most frequently presented to teachers in justification for what diagnostic specialists thought should be done with or to individual students.

Today, classroom teachers are called upon to be much more actively involved in student assessment and evaluation. Their participation is important to the interdisciplinary evaluation process, and their ability to assess student skills and progress is vital to carrying out the Individualized Educational Plan (IEP).

To carry out these responsibilities on the one hand, teachers need to know much more about issues and assumptions involved in traditional methods of psychoeducational assessment. They need to be familiar with both the uses and abuses of diagnostic assessments and they need to be sensitive to the biases which often enter into these assessments. In effect, they must be enlightened consumers and monitors of this process.

On the other hand, teachers need to be better able to participate actively in assessments which focus less on student characteristics and more on student skills, student needs and the instructional environment. It is in this area that Public Law 94-142 most challenges those involved in the assessment process. The principle that student assessment should lead to an individualized program appropriate to the needs of the individual child requires far more than the categorization of children through diagnostic procedures. On the contrary, Public Law 94-142 requires that a student's present skills be known and progress toward reasonable goals be monitored. This requirement in the assessment process distinctly falls in the realm of good teaching.

The materials presented here are designed to provide faculty members of colleges of education with information and activities to help prepare regular class teachers to undertake their vital role in the assessment and evaluation process. The contents focus on what teachers should know about standardized psychometric evaluations and also on assessment procedures which can and ought to be carried out by classroom teachers themselves. Since this module only outlines a few of the most important issues regarding assessment and evaluation procedures, references have been listed in the back of the module that elaborate on these issues.

Other modules in the total set are related to this topic. The following should be noted in particular:

9A: Writing goals, objectives, and IEPs;
10A: Educating handicapped children: Judicial and legislative influences;
8A: Formal observation of students' social behavior;
4A: The functions and responsibilities of support personnel.

Contents

Within this module are the following components:

Page

Set of Objectives. The objectives focus on the teacher educator rather than on a student (preservice teacher). The objectives identify what can be expected as a result of working through the materials. The objectives which apply to teachers are also identified. They are statements about skills, knowledge and attitudes which should be part of the "common body of practice" of all teachers. 5

Rating Scales. Scales are included by which a teacher educator could, in a cursory way, assess the degree to which the knowledge and practices identified in this module are prevalent in the existing teacher-training program. The rating scales also provide a catalyst for further thinking in each area. 6

Self-Assessment. Specific test items were developed to determine a user's working knowledge of the major concepts and principles in each subtopic. The self-assessment may be used as a pre-assessment to determine whether one would find it worthwhile to go through the module, or as a self-check after the materials have been worked through. The self-assessment items also can serve as examples of mastery test questions for students. 7

Rationale and Knowledge Base. The brief statement summarizes the knowledge base and empirical support for the selected topics on pupil assessment. The more salient concepts and strategies are reviewed. A few brief simulations/activities and questions have been integrated with the rationale and knowledge base. 10

Page

Bibliography. A partial bibliography of important books, articles and materials is included after the list of references. 80

Articles. Five brief articles (reproduced with authors' permission) accompany the aforementioned components. The articles support and expand on the knowledge base. 82

Objectives

On completing this module, you will be able to identify appropriate and inappropriate procedures for assessing students with consideration of the following conceptual, methodological and ethical issues:

1. The testing debate;
2. The educational role of professionals and Public Law 94-142;
3. Assessment methodology;
4. Relevant assessment variables;
5. Assessment procedures;
6. Norm-referenced and criterion-referenced assessment;
7. The relationship between assessment and instruction;
8. Evaluation of intervention effects;
9. Interpretation and communication of assessment information.

Reasonable Objectives for Teacher Education

Students should have knowledge, practical skills and commitment to professional performance in the following areas relating to assessment and evaluation procedures:

1. Methods and underlying assumptions, issues and biases of norm-referenced psychoeducational assessment.
2. Appropriate and inappropriate uses of norm-referenced and criterion-referenced assessment.
3. Administering and interpreting norm-referenced and criterion-referenced assessment instruments.
4. Developing criterion-referenced assessment procedures directly relevant to the curriculum of a classroom.

Rating Scale for the Teacher Preparation Program

Check the statement that best describes the level of your present teacher education program on the topic of assessment and evaluation procedures:

____1. Students being prepared for teaching have been introduced to educational and psychological assessment, but have little understanding of the issues and assumptions in traditional psychometrics. They see their own role in student assessment and evaluation primarily as giving grades. They see other forms of assessment and evaluation as being in the realm of specialists.

____2. Students being prepared for teaching have received instruction on the issues, assumptions and methods of standardized testing. They have learned how standardized instruments are developed and how results are to be interpreted. They perceive standardized normative assessment procedures as adequate measures of student abilities.

____3. Students being prepared for teaching are taught the issues and assumptions underlying both norm-referenced and criterion-referenced assessment. They perceive both as useful when used appropriately, but have little or no experience in actual assessments.

____4. Students being prepared for teaching are taught the issues and assumptions underlying both norm-referenced and criterion-referenced assessment. Students are given practical experience using and interpreting norm-referenced and criterion-referenced instruments.

____5. Students being prepared for teaching are taught the issues and assumptions underlying both norm-referenced and criterion-referenced assessment. Students have practical experience using and interpreting commercially developed norm-referenced and criterion-referenced assessment based on the curriculum of a particular class.

Self-assessment

The following items are intended to assess your understanding of the significant issues in educational assessment.

Short Answer Items

1. Define the purpose of educational assessment.
2. What is the difference between testing and assessment?
3. Outline the requirements of Public Law 94-142 as they relate to the educational assessment of minority students.
4. Identify two sources of information about a student's functioning in educational settings.
5. Identify two factors to be considered when making informed decisions about a student's educational functioning.
6. Identify the major components of a systematic assessment procedure.
7. Identify two sources of inconsistency or error in methods of gathering assessment information about a student.
8. Describe the nature of the information that can be gathered from achievement tests.
9. Identify two factors that can negatively influence the validity of a test.
10. What are some purposes of educational assessment?
11. What is the difference in the nature of information obtained from a norm-referenced test and a criterion-referenced test?
12. Describe domain-referenced testing.
13. What are five major considerations in setting a criterion for a student's performance?

14. State two areas which are of concern when evaluating treatment effectiveness.

15. Define treatment validity.

16. What are the three mandates laid down in Public Law 94-142 that relate to the interpretation and communication of assessment information?

In front of the next 11 items, place a "T" if you think the statement is true, or an "F" if you think the statement is false.

____1. Tests are either "good" or "bad." Poor programming is the result of using "bad" tests.

____2. The purpose of assessment is to prescribe a fixed treatment that continues until a cure is achieved.

____3. The requirements of Public Law 94-142 necessitate a comprehensive approach to assessment team membership and to instrumentation.

____4. School psychologists can be most useful by confining their activities to the administration of standardized tests of ability and achievement.

____5. When placement decisions for a student are made, information from norm-referenced tests provides the most comprehensive data-base.

____6. Information obtained from a student's parents can assist in verifying information given by the referring teacher.

____7. Students' achievement test scores reflect both their ability and the instructional effectiveness of their educational program.

____8. Human error is a minor source of inconsistency in the assessment process and compensations for unreliability can easily be made.

____9. It is the responsibility of the test developer to ensure test validity. Once this has been demonstrated, the professional need have no concerns about using the test.

____10. An advantage in using norm-referenced tests is their adaptability to a specific curriculum sequence.

____11. To be of maximum use to the teacher, criterion-referenced testing must be integrated into the day-by-day functioning of the classroom and must not be separated out as a "testing" activity.

The Testing Debate

The role of psychometric testing within the educational system has long been debated. A view of testing as a sorting procedure which results in students being divided into those who will succeed and those who will not is common. Testing is sometimes considered a tool for social decision-making about the futures of millions of children (Burton, 1978). Such decision-making, in too many cases, has resulted in inappropriate placement of children in educational programs and curricula, with life-long deleterious effects. It has also been noted that inappropriate decisions seem to occur more frequently among certain minority groups, and this has been seen, at least in part, as a result of placement decisions based on culturally biased psychological tests (Larry P. vs. Riles, 1979; Mercer, 1973).

In the testing debate, antagonists cite not only political and social abuses of testing (Kamin, 1975), but also testing devices that are technically inadequate, are normed on populations which bear no resemblance to the pupils being tested, or inadequately sample the behaviors that the instrument supposedly tests (Salvia & Ysseldyke, 1978). Adelman and Taylor (1979) present several conceptual, methodological and ethical assessment issues which merit discussion.

Key Conceptual, Methodological, and Ethical Assessment Issues

A. Conceptual

Current practices reflect a rather inadequate appreciation of the following matters:

1. Whose interests are and should be served by the assessment activity (e.g., the society, the client, the intervener)?
2. Who sets criteria determining assessee status, needs, problems, and progress?

3. Who decides on whose interests should be served and on who is to set criteria?
4. What are the various tasks which currently are subsumed under the term "assessment diagnosis"?
5. What are the limitations in perspective of the models of cause and correction upon which current activity is based?
6. What are the assessor biasing factors which need to be systematically accounted for in assessment activity?
7. What are the assessee motivational and developmental factors which need to be systematically accounted for in assessment activity?

B. Methodological (technical, practical)

While widely acknowledged, the following problems continue to place major limitations on assessment activity:

1. The more complex the assessment objectives, the lower the reliability of the total set of measurement procedures.
2. Construct validity often has not been demonstrated satisfactorily by scientific standards, e.g., the "validation" procedures, when undertaken, often are tautological.
3. Predictive and postdictive validity appears to diminish, in some instances at an exponential rate, the more distant in time the assessment data being gathered are from the criterion being predicted to (or from the original factor causing the behavior under investigation).
4. There is a sparsity of systematically gathered and agreed-upon norms and standards for making judgments ("good-bad," "normal-abnormal," "success-failure") thereby resulting in idiosyncratic variations on judgments which are often beyond accountability.

5. The utility of a procedure may be judged as much (or more) on the basis of its marketability and the current absence of a feasible alternative, as on its efficacy (e.g., its validity with regard to the decisions being made, its ability to add information beyond base rate levels).
6. The costs of assessment practices are escalating; time demands often are extensive; referral practices tend to overrely on "old boy" networks, etc.

c. Ethical

Ethical practices reflect inadequate efforts to systematically detect and minimize the following iatrogenic effects:

1. Misidentification of the cause of a problem (e.g., false negatives and positives such as those resulting from over-reliance on person-focused tests and observer-assessor biases).
2. Misprescriptions related to subsequent intervention procedures.
3. Violations of rights (e.g., failure to get truly informed consent, invasion of privacy, denial of access to assessment reports and of the right to correct the record).
4. Negative repercussions of assessment processes or products (e.g., increasing feelings of anxiety, incompetency, and lack of self-determination, increasing over-reliance and dependency on professionals, initiating self-fulfilling prophecies and stigmatizing effects).
5. Failure to pursue professional responsibility with regard to improving standards of practice and advancing knowledge (including collusion with an inadequate status quo).

Note. From "Initial psychoeducational assessment and related consultation," by H. S. Adelman and L. Taylor, Learning Disability Quarterly, 1979, 2, 52-64.

Clarification of the Issues

Consideration of the following three points may help clarify these testing issues:

1. The concentration on the abuses that have occurred in psychoeducational assessment may blur the distinction between tests and the decisions made from test data. Many abuses occur because of ignorance or overzealousness on the part of the diagnostician or decision-maker. Tests are not necessarily "good" or "bad"; such value judgments can be reserved for evaluating how test results are used in screening, placement, program planning, and evaluation of individual pupil progress.

2. Testing is but one method of assessing a child and is only one tool to use in making educational decisions. Educational personnel who fail to make use of such tecnniques as the interview, observation, and environmental analysis, as well as their own professional judgment have gathered inadequate information to make sound educational decisions (Shertzer & Linden, 1979).

3. "The process of assessment is a constructive...flexible, continuous process, leading not to a fixed prescription of treatment until a cure is achieved, but to an ongoing program which may frequently be modified in the interests of the student's life situation and of a reduction in his current specific difficulties." (Clarke & Clarke, 1975)

The Educational Role of Professionals and PL 94-142

Individual roles and responsibilities of professionals in educational settings are, of course, largely determined by the nature of their professional training. This factor, as well as the institutionalization of specialized roles, has often led to a compartmentalization of responsibilities. In effect, this compartmentalization has created a view of the child as a group of separate and unique entities. But the practices which have resulted from such narrow perspectives are no longer feasible under Public Law 94-142, which requires that a student's individual education program be "developed by a team consisting of the child's teacher, a person other than a teacher who is qualified to provide or supervise the provision of special education, one or both parents, the child (when appropriate), and other persons who are brought in at the discretion of the school." The purpose of the IEP is to provide an overall program of special education and related services and it should include a justification for those services and placement. In addition, the IEP must include the objective criteria that will be used to evaluate the child's achievement. A multidisciplinary assessment team and multiple assessment methods are necessary.

Traditional Psychoeducational Assessment

Traditionally, roles involving the assessment of children have often restricted school personnel to the administration of standardized tests of ability and achievement. Information from these tests was used to predict a student's success or failure during school and in later life. Diagnosis of a student's abilities or deficits was often confined to norm-referenced statements based on composite scores and often had few direct implications for instructional intervention. Such diagnostic practices were enlisted to support special education practices of classification according to handicap for the purposes of placement, and for making claims for federal funding, but told little about what type of educational plans should be designed for students diagnosed as evidencing any one of several handicapping conditions. But these traditional assessment practices are rapidly changing.

Reynolds (1975) has called attention to changes in special education that are also effecting changes in assessment practices. He states:

> We are in a zero-demission era: consequently, schools require a decision orientation other than simple prediction; they need one that is oriented to individual rather than institutional payoff. In today's context the measurement technologies ought to become integral parts of instruction designed to make a difference in the lives of children and not just a prediction about their lives. (p. 15)

Cognitive Entry Characteristics

In selecting assessment procedures that will make a difference, the following observation regarding "cognitive entry characteristics" is appropriate:

> Quite in contrast to intelligence and aptitude indices are cognitive entry characteristics. These are the specific knowledge, abilities, or skills that are essential for the learning of a particular school or a particular learning task. Such prerequisites typically correlate +.70 or higher, with measures of achievement in a subject. Furthermore, when they are identified and measured, they replace intelligence and aptitude tests in the prediction of later achievement. That is, intelligence or aptitude tests add little or nothing to cognitive entry measures for the prediction of school learning. All of this is to say that cognitive entry characteristics have a high relationship to achievement, and they have an obvious causal effect on later achievement. This is especially true where sequential learning tasks are involved where it may be impossible to learn task B without prior adequate learning of task A. (Bloom, 1980)

This distinction between intelligence and aptitude tests and specific knowledge and skill tests may be very important for special education populations where the predictive value of intelligence and aptitude tests is lessened by the variability within handicapped groups and the small percentage of handicapped individuals included in population samples used to develop test norms.

The comparative predictive value of different types of test information was graphically demonstrated in a study by Hofmeister and Espeseth (1970). This study investigated the effectiveness of measures of mental age (Stanford-

Binet), language age (Illinois Test of Psycholinguistic Abilities) and reading achievement (Wide-Range Achievement Test) as predictors of progress with a population of moderately retarded youth in a selected reading program. The predictive coefficients were: -.11 for mental age; +.11 for language age, and +.78 for the reading achievement test. The discriminating item in the reading test responsible for identifying successful readers was the letter naming item. In essence then, a two-minute test of a student's skill in attaching names to letters was found to be far more useful for assessing a student's ability to participate in a specific reading program than the estimates of mental age or language age. Hofmeister and Espeseth concluded that:

> ...variables such as MA and LA, which are often thought of as psychological correlates of reading ability, were comparatively ineffective as predictors of reading achievement with the group observed in this study. The results also suggested that it might be more profitable to look at task oriented variables which are concerned with determining the position of the individual in the learning sequence. (p. 107)

PL 94-142 and Nondiscriminatory Evaluation

In order to ensure appropriate educational placement, equality of educational opportunity, and the right to ethnic dignity and respect, and to prevent unfair stigmatizing of students, the Education for All Handicapped Children Act of 1975 (PL 94-142) mandates the following assessment procedures:

1. The testing and evaluation materials and procedures will be selected and administered so as not to be culturally discriminatory.
2. Such materials and procedures are to be provided in the child's native language or mode of communication.
3. No single procedure or test can be the sole criterion for determining the appropriate education program for the child (Section 612).

Under PL 94-142, it is the pupil needs and not a district-wide testing program that determines the instrument and associated assessment processes. The following examples sample some of the diversity of assessment practices possible under 94-142.

1. You have to assess a 19-year old in a vocational program. Suggestion: You might review "The Assessment of Applied Academic and Social Skills," (Forness, Thornton, & Horton, 1981). See this article in the appendix.
2. A parent requests an evaluation of her 17-year old son's writing skills. Suggestion: James A. Poteet's "Checklist of Written Expression" would be a practical instrument that would provide a series of specific short-term objectives for an IEP goal in this area. See this article in the appendix.
3. The parent of a severely handicapped child expresses interest in establishing an IEP goal on leisure skills. Suggestion: Wehman and Schleien's (1980) suggestions for the "Assessment and Selection of Leisure Skills for Severely Handicapped Individuals" would be appropriate for this assessment assignment. See this article in the appendix.

Assessment Methodology

Educational assessment is a multifaceted and systematic process carried out for the purpose of making decisions about the performance of students in their current school setting. Assessment is more than testing. It is the systematic process of using information from all possible sources in order to make educational decisions about students. A systematic approach to educational assessment is one in which a variety of assessment methods are used and in which all relevant assessment variables are considered.

Assessment Methods

Information about a child's functioning in educational settings may be gathered through:

1. Inspection of the child's record files. Information reported in cumulative and other available records can be helpful in determining if there are factors that might account for problems the child is having, if there are trends in the growth of the problem, and if there are other relevant factors that need to be evaluated.

2. Informal consultation with others who know the child. Consultation with persons who know the child well (e.g., peers, siblings, parents) will lead to a more comprehensive assessment.

3. Structured interviews. Interviews with the parent, the student and school personnel can yield information about the student's areas of difficulty and determine what resources have been previously employed in working with the student.

4. Observation. Through observation, assessment data is collected on the behavior of the student as it occurs naturally in the environment. Such data can focus on specific facets of behavior (e.g., student-teacher interaction), can indicate areas in need of further evaluation, can confirm other information gleaned from records, and can provide insights that may lead to an improved program for the student. Observations may be formal or informal and include anecdotal records, interaction analysis, checklists, and rating scales. Observation procedures can be both long-term and short-term processes.

5. Norm-referenced tests. Standardized tests can be used for both screening and diagnostic purposes to determine how one child compares with other children.

6. Criterion-referenced tests. This type of test measures a child's level of skill development in a particular area. It is especially useful for planning purposes because criterion statements can be used as goals and short-term objectives for instruction.

Relevant Assessment Variables

Educational decision-makers should gather educational information about the "whole" child. They should therefore consider all relevant factors which, because of their interactions with the child, may influence the child's functioning. These factors will vary with the child and the particular educational problem, but should include:

1. Classroom environment. The classroom environment refers to the physical arrangement of objects within the classroom (seating arrangements, materials, learning centers, etc.) which have the potential to impact learning, social interaction patterns, time on-task, interests, achievement levels, etc.

2. Curriculum. A thorough knowledge of the curriculum presented to the child is necessary when making decisions about that child's educational functioning. Assessors should be aware of:

   a. conceptual level and sequence of student activities;
   b. number of academic tasks and task sequences;
   c. evaluation methods used in classroom;
   d. variety and levels of instructional materials;
   e. appropriateness of curriculum and materials for the child.

3. Instructional behavior. Teacher behavior is the unifying factor which brings the environment and the curriculum together to impact on the child.

To be effective, assessment processes must be conducted so that specific remedies expressed in changes in instructional behavior can be generated.

Engelmann, Granzin and Severson (1979) observed that

> ...the teacher achieves behavioral change only by manipulating environmental events. This point is extremely important. It follows that the remedy must clearly imply manipulation of those environmental events. It must tell teachers what they are doing wrong and what types of different teaching behaviors they should implement. The remedy must be specific and concrete because teaching always involves specific and concrete acts. (p. 356)

When making decisions about the functioning of students in educational settings, it is important to consider the three factors listed above. However, the primary focus during the assessment process should be on the pupil's problems and remedies that can be implemented with the available resources.

Several models have been developed for systematic educational assessment. Underlying all models is the movement of decision-makers from information which is general, broad, and only assumed to be accurate, to information which is specific, precise, and valid. The purpose of gaining valid information through the assessment procedure is twofold. First, professionals must be in a position to make valid intervention decisions about a pupil with special needs. Second, the assessment information obtained must provide the basis for an evaluation of the success of any interventions that are implemented.

In the following section, an example of a systematic assessment model is presented. Assessment methods and sources of information about the child are integrated into an assessment procedure. The procedure outlined is one which might be followed in an interdisciplinary evaluation process. However, it certainly is the case that many of the procedures described may be used to advantage by school personnel outside the formal assessment process.

A Model for Systematic Assessment

Screening

Referral and verification. The referral process is the first step in assuring the appropriate programming of children with suspected special educational needs. Referrals are typically initiated by counselors, psychologists, classroom teachers or resource teachers, but may be initiated by others, including the students themselves, or their parents. Information leading to the verification of educational problems for which a student has been referred can come from examination of the student's cumulative records and direct observation of the student's behaviors. Data on the student's personal, medical and educational history from cumulative records can contribute to referral information and allow educational personnel to discern the pattern of events leading up to the present situation. Observations can provide comprehensive, detailed, and specific information about the behavior of the student and about the contexts or environments in which the observations are made. Much of the direct observation can be carried out by teachers. In fact, data collected by teachers before the formal assessment process is initiated often suggests possible solutions to the present problem, or that the student's behavior does not substantially deviate from that of other class members (see "Formal Observation of Students' Social Behavior" by F. Wood, in this series).

In observing the student in interactions, the focus should be on the identification of habitually-occurring events which reinforce the behavior being studied. Both social and nonsocial consequences which maintain the behavior must be identified, and an estimate of the student's responses to such stimuli should be made. It should also be possible to identify means of structuring the environment in the future. Once educational problems and target behaviors for the student have been identified, environmental contingencies which will reinforce the target behaviors can be determined.

One method of collecting the information outlined above is through a behavioral interview. While the behavioral interview is probably one of the least structured assessment strategies, it can nevertheless support the more objective data gathered through systematic observation. The following outline is suggested to ensure that all possible information is collected from the individual and that data relevant to the formulation of an achievement plan is gathered.

1. An initial analysis of the problem situation should be made, in which problematic behavioral excesses and deficits as well as nonproblematic behavioral assets are specified.

2. The problem situation should be clarified by identifying the individuals who object to the problem behavior and who may be affected by any behavior change made by the student. The clarification should also help to specify the conditions under which the behavior occurs.

3. A motivational analysis should be carried out in which reinforcers (both positive and negative) that may be maintaining the problem behaviors, or that may be useful in shaping more appropriate behaviors, are specified.

4. A developmental analysis should be made in which several questions are raised about the biological, sociological, and behavioral changes that may be pertinent to the problem behavior.

5. An analysis of the student's self-control should be completed, in which the limitations, conditions, and methods of self-control are defined.

6. Social relationships should be identified and the influence of significant others on the student specified.

7. The student's social, cultural, and physical environments should be examined with consideration of cultural norms relating to the problem behavior, the similarity of norms in different settings, and various environmental restraints impinging on the student.

Schertzer and Linden (1979) recommend that information be obtained from significant others in the students' lives, e.g., parents and school personnel. These second-party reports help in obtaining information about the student's likes and dislikes (objects, events, places, activities, people, etc.), and in identifying the perceptions of school personnel and parents as to the nature of the problem they see with the child. In addition, assessment instruments which structure the perceptions of these significant others, such as the American Association on Mental Deficiency Adaptive Behavior Scale (AAMD, 1974) or the Devereaux Elementary School Behavior Profile (Spivack & Swift, 1967), can assist in the gathering of second-party report information.

Appropriate screening procedures also require that information from previous assessments of the student's performance on standardized tests of general verbal and nonverbal ability (such as intelligence tests) and standardized educational achievement tests (such as reading, spelling, and math) be examined.

Standardized achievement test data. Achievement tests are devices that assess a student's skill development in academic content areas. Most tests are commercial instruments that sample the products of assumed past formal and informal educational experiences and measure the extent to which an individual has profited from schooling and/or life experiences as compared to others of the same age or grade. The scores assigned to a pupil reflect both pupil ability and the instructional effectiveness of the educational program. When using information from achievement tests, the assessor must be aware (a) of the nature of behaviors sampled by the test, particularly the relationship of specific curricula in use in the school district; (b) what specific items the student has passed and failed, and consistent failure patterns; and (c) that in the screening procedure, achievement tests determine, in a global way, only the child's current level of functioning.

Standardized achievement tests, such as the Stanford Achievement Test and the Iowa Test of Basic Skills are often used as screening devices. These tests in fact measure the extent to which a student has benefited from past schooling, compared to others of the same age or grade level. Knowing the nature of the subtests in these batteries allows some information to be obtained about the remediation needed by individual students and provides a general idea about where to start additional diagnostic assessment.

Evaluation of Screening Data and Planning Interventions

In this step, school personnel evaluate the screening data and their reliability. It is important to be aware of the factors influencing reliability and validity of observation and test information.

Information gained from interviews will vary among interviewees, and these inconsistencies mean that all interpretations must be viewed as tentative hypotheses to be verified or refuted with further evidence. For example, interview data from a teacher with limited observational skills who has not systematically collected data, or is hoping to have a difficult child removed to a separate special education class, will often be unreliable and biased and may be of little use in the assessment process.

Observational tools must be used with care and precision. Misinterpretation and misuse of observational techniques will distort estimates of individual behavior patterns which are at best only samples of behavior. Interpretation of information must include attention to potential problems of sampling error and of inadequate sampling. Observer bias and the possibility of the individual's awareness of being observed, in conjunction with the variability of behavior and the selectivity of observation, are threats to the validity and reliability of observational tools and techniques.

The direct observation of the pupil in instructional settings is one way to correct two common weaknesses in assessment practices, namely, the lack of emphasis on alterable variables and the implicit assumption that the child is always the problem. Too often assessment practices allocate extensive resources to collecting data on variables that cannot be manipulated to help the pupil. This might occur by collecting extensive data on community and home variables that cannot be altered with the resources available to the school. A similar problem occurs with the assessment of some school-related behaviors. For example, the assessment of some categories of perceptual-motor abilities presents a problem because the etiological relationship with academic skills is unclear and the research suggests that our procedures for "training" these abilities are generally ineffective (Kavale & Mattson, 1980).

The direct observation of pupils in instructional settings allows for the collection of information on instructor behavior which may be an important part of the problem. Variables such as instructional methods, instructional materials, classroom organization, teacher feedback and reinforcement practices are all important variables that can be changed.

Generalizability of Observational Data

Too often observational data lacks validity because of incomplete sampling practices. Collection of observational data in high school programs presents obvious problems because of the fragmentation of curriculum and instructional supervision. Data collected in one subject matter area may have limited value in making generalizations to other subject matter areas and other teachers.

Even with the same teacher, several observations across several lessons may be necessary to achieve acceptable levels of generalizability. Capie (1981) found that in observing students' on-task behavior, 20 observations in each of five lessons were necessary to achieve a generalizability coefficient of .80 (a generally accepted level of instrument reliability). Capie noted that to ensure reliability of the observational samples "the number of lessons observed is far more important than the amount of observations within lessons." (p. 18)

Factors affecting test validity. Whether to use a test for a certain purpose and how to interpret the information yielded by the test are decisions which should be governed by the validity of the test. Validity is specific both to purpose and population. Despite efforts of test developers to ensure content, construct, and predictive validity, tests are only valid for certain purposes. Selection of standardized tests must be made with the purpose of the test in mind. These questions should be considered:

1. Was the test designed to be used for screening purposes?
2. Was the test designed to be used to place a student in a specific curriculum?
3. Was the test designed to be used to assist teachers and administrators in planning individual or group educational programs?
4. Was the test designed to be used for program evaluation?
5. Was the test designed to be used to assess individual progress?

On another level, and related to the question of validity, are questions that must be asked to determine the accuracy of information yielded by the test:

1. Was the test technically adequate, that is, did it specify adequate information about test administration, standardization, reliability and validity?

2. Assuming that the test was technically adequate, was it used for the right purpose? For example, was a measure of receptive vocabulary, such as the Peabody Vocabulary Test, used as a measure of intelligence?

3. Assuming that the test was technically adequate, was it appropriate to the child? Did the student come from a cultural group comparable to those in the norming sample? Was the test age-appropriate? Did the child possess physical and mental characteristics comparable to those in the norming sample? Did the test require responses that a child with motor, visual, or hearing impairments would be unable to perform? Was the tester aware of any medication the child was taking? Many drugs, such as Mellaril, Dilantin, Valium, or Ritalin may affect test results by influencing the nature and rate of responses (Wysocki, Fuqua, Davis, & Breuning, 1981).

4. Is the assumption being made that norm-referenced scores such as grade and age equivalents, or composite scores such as IQ or Social Quotient (SQ), are giving adequate representations of skills and deficiencies?

5. What information does the tester have about the child's background and current status and performance that will assist in accurate interpretation of test scores?

6. What relationship exists between the test and the curriculum in which the student is currently placed? Is the test measuring knowledge (and the chance to acquire it) or is it measuring variables that may not be in the student's repertoire of learned behavior?

If these questions cannot be responded to adequately, the test information must be interpreted with caution. In such cases, more appropriate assessment is necessary to provide an accurate picture of the student's abilities and deficiencies so that an appropriate and relevant program can be developed.

Intervention

The purpose of an educational assessment is to provide: (1) recommendations for instructional and behavioral goals and objectives, (2) recommendations for educational placement, (3) measures of students' present skills, (4) measures of students' progress toward established goals and objectives, and (5) evaluations of the effectiveness of students' educational programs. In the following section, some methods for using assessment information for these purposes will be presented.

Activities

Investigate the assessment procedures used in a school. To what extent are decisions based on standardized and achievement tests? To what extent are data related to a student's actual behavior and specific skills used in the assessment?

Plan to spend 30 minutes with a child. For 15 minutes ask the child questions from a standard psychological test (e.g., WISC-R, Stanford-Binet, ITPA), and for the other 15 minutes have the student answer problems from a personally constructed, sequenced list of arithmetic problems. Compare the usefulness of the two assessment procedures in making statements about the child's educational abilities and needs.

Norm-Referenced and Criterion-Referenced Assessment

Quality educational programming requires that assessment instruments be used to ensure that each pupil enters a curriculum with the necessary prerequisite skills. A testing technology has been developed that is concerned with defining the relationship between an individual's skills and instructional sequences. Glaser (1963) in an article on the measurement of learning, defined two types of tests. One type, the norm-referenced test, evaluates the child's performance by comparison with that of the other children. The intelligence test is a classic norm-referenced test. The other type, the criterion-referenced test, evaluates the child in terms of some learning task at some absolute standard that is independent of other children's performances. The criterion-referenced test is designed to identify the relationship between a pupil and a specific instructional task. If a teacher wants to know what types of number tasks a child has mastered, then a criterion-referenced test would generally be called for.

Although a test is usually constructed as either norm-referenced or criterion-referenced, a person may use one test for both purposes. If we were to examine the items on a standardized achievement test for patterns of successes or failures on specific instructional tasks, we would be using it as a criterion-referenced test. If we used a test developed as criterion-referenced to rank the pupils in a class, we would be using it as a norm-referenced test because the child's score (rank) would describe him in terms of his relationship with others (the rest of the pupils in the class).

Normative Testing and Instruction

The need to develop Individualized Education Programs (IEPs) for special education pupils requires reconsideration of the role of norm-referenced testing procedures. Normative tests are typically designed for such purposes as prediction, selection, and comparison. Special education has a long history of attempting to adapt these tests for individualized instruction. However, except for a few standardized diagnostic tests which have both norm-referenced and criterion-referenced properties, the majority of norm-referenced tests provide limited information for guiding the teaching of specific skills. A major problem encountered in the use of norm-referenced tests is that most are not designed to relate directly to a specific curriculum sequence. The discrepancies between norm-referenced and criterion-referenced test information become especially important when using such information for individual program planning.

Two Types of Testing Compared

Norm-Referenced	*Criterion-Referenced*
• *Evaluates individual performance in comparison to a group of persons.*	• *Evaluates individual performance in relation to a fixed standard.*
• *Is used to evaluate a student as "below grade level," "at grade level," or "above grade level."*	• *Not concerned with grade level descriptions.*
• *Fails to indicate where students have mastered the spectrum of instructional objectives.*	• *Identifies the individuals who have mastered the spectrum of instructional objectives.*
• *Generally poor aides in planning specific instructional programs.*	• *Geared to provide information to be used in planning instruction.*
• *Is often vague in relation to instructional content.*	• *Is content specific.*
• *Does not operationally define mastery and/or success.*	• *Operationally defines mastery and/or success.*
• *Applies poorly to the individualization of instruction.*	• *Applies directly to the individualization of instruction.*

• *Is not concerned with task analysis.*	• *Depends upon task analysis.*
• *Test results interpreted in reference to a person's position in relation to the scores of others.*	• *Test results interpreted in reference to a person's position in relation to the curriculum.*

--adapted from Housden, J. L., & LeGear, L. An emerging model: Criterion-Referenced evaluation. *Thrust, April 1973.*

In the example in Table 1, data were collected from twenty-four pupils referred for special education services as spelling failures. Two types of test data were collected: (a) norm-referenced data from a standardized test, and (b) data from a criterion-referenced test that identified the placement of a child with regard to the specific spelling curriculum in use in the school. Generally, a criterion-referenced test is not used to identify grade levels because this usually entails moving from specific practical information to more general and less prescriptive information. For purposes of comparison in this case, however, a grade level score was assigned to the criterion-referenced data. This grade level score indicated the level in the curriculum set for that grade at which instruction should begin. An observation of the scores in Table 1 reveals that the difference between the overall mean grade level scores on the two tests was not great (0.14). However, when the pupils are considered as individuals, relationships between normative scores (standardized achievement test) and criterion tests (curriculum placement test) are so great that the rather common practice of using standardized achievement test data for individual program planning must be questioned. In analyzing the content of the two tests, it was found that the criterion-referenced test content was consistent with the specific spelling program in use in the school, which placed a strong emphasis on controlled sequences built around the phonetic structures of words. In this

Table 1

SPELLING FAILURE REFERRALS

Norm-Referenced Score--Curriculum Placement

	Ach. Test Score* (Norm-referenced)	Curr. Placement* (Crit.-referenced)	Differences in Placement*	
Grade 3 N = 6	1.1	2.0	- .9	X = -.45**
	2.0	2.2	- .2	
	2.0	2.2	- .2	
	1.7	2.0	- .3	
	1.7	2.0	- .3	
	1.4	2.2	- .8	
Grade 4 N = 6	1.8	2.2	- .4	X = -.23**
	1.8	2.0	- .2	
	2.1	2.2	- .1	
	1.5	2.2	- .7	
	1.5	2.0	- .7	
	2.9	2.2	+ .7	
Grade 5 N = 6	3.2	2.8	+ .4	X = .55**
	3.2	2.2	+1.0	
	3.2	2.2	+1.0	
	3.4	2.2	+1.2	
	3.2	2.8	+ .4	
	1.9	2.6	- .7	
Grade 6 N = 6	4.5	3.4	+1.1	X = .68**
	2.9	2.8	+ .1	
	2.9	3.4	- .5	
	4.0	2.8	+1.2	
	4.8	2.6	+2.2	
	2.9	2.8	+ .1	

Total number of subjects = 24
Difference between means = .14
Correlation between normative and criterion-referenced test scores = .65

*All scores expressed in grade levels
**Positive score indicates normative test higher

instructional program, irregular words were emphasized much later than they are in some other types of spelling programs that use different theoretical rationales such as "frequency of use." The content of the standardized achievement test stressed irregular words much earlier than did the criterion-referenced test; it was, therefore, of little value for making decisions related to the phonetically oriented curriculum being used for the twenty-four children listed in Table 1. Thus, it is important to choose tests that relate to the curriculum being used in order to obtain accurate data in developing the child's individual program.

The Process of Diagnosis

> The process of diagnosis in an instructional setting has been likened by some to the medical process from which the term has its roots (i.e., a doctor examines the patient's symptoms and the pattern of these symptoms leads him to the diagnosis that a certain disease of dysfunction is present).
>
> This analogy is spurious, however, for it implies that the learning diagnosis identifies or defines a knowledge "disease" or a disorder rather than a specific need or a cluster of needs, which is what actually takes place. Diagnosis in education, then, is actually a needs assessment process.
>
> The instructional manager uses his professional skills and the instruments and techniques available to him to examine the present state of the learner in relation to pre-established learning outcomes or objectives, usually arrayed in a continuum or hierarchy of progression (Hickey & Hoffman, 1973, p. 35).

Diagnostic Devices

There are a number of standardized diagnostic tests available that have both norm-referenced and criterion-referenced properties. Many standardized diagnostic tests can tell the teacher whether a child is above or below average for his chronological age in the area tested, and in sub-areas that make up the general area. These tests also have some value for program evaluation, since they demonstrate student achievement particularly if accountability data are required in simple, readily understood terms. Their weakness in this regard is, of course, that student achievement is directly tied to the content of instruction and measured progress is highly influenced by the congruence between test and curriculum.

In many settings where quality student assessment procedures exist, broad standardized tests are used in combination with criterion-referenced tests that are more specific to the local curriculum and associated materials. A standardized diagnostic math test might indicate that a child is weak in addition combinations, but a more specific instrument will be needed to determine which addition combinations need to be taught to the child.

The skills assessed by diagnostic reading tests include those of oral reading, comprehension, word attack, word recognition, and rate of reading. Tests commonly used are the Gray Oral Reading Test, Gates-McKillop Reading Diagnostic Tests, Durrell Analysis of Reading Difficulty, Stanford Diagnostic Reading Test, and the Woodcock Reading Mastery Test. Another approach is exemplified by the Fountain Valley Support System in Reading, a criterion-referenced testing system. The Fountain Valley system includes pupil profiles which provide a record of individual pupil achievement and facilitates the monitoring of pupil progress.

Diagnostic assessment in mathematics leads to specific information about a student's performance in content areas (e.g., numeration, fractions, algebra) operations, and applications (e.g., measurement, time, problem-solving, money). Three most commonly used diagnostic mathematics tests are the Key Math Diagnostic Arithmetic Test, the Stanford Diagnostic Mathematics Test and Diagnosis: An Instructional Aid in Mathematics.

The Relationship of Domain-Referenced Testing to Criterion-Referenced Tests

The quality of a criterion-referenced test depends on the degree to which a skill area is clearly identified and represented within the test. Some criterion-referenced tests are poorly constructed. In recognition of this fact, Hively (1974) coined the term "domain-referenced." A domain-referenced test is one in which the emphasis is placed on precisely identifying skill areas (domains). The test items of a domain-referenced test are selected so that one can be certain that a child who meets criteria on the items which represent a domain would be able to master all other possible items of the domain when they were encountered. For example, if a series of test items was prepared to test a child's mastery of a domain, such as long division of decimal fractions using single digit divisors, the test items would have to include examples with zero in the dividend. Without such examples involving zero, there would be little quarantee that performance on the test items would generalize to all the types of long division examples a child might encounter.

In describing the term "domain-referenced," Donlon (1975) noted:

> The label "criterion-referenced" has what Hively calls "surplus associations." Further, Hively has recognized that our educational objectives are seldom detailed: We say we want the child to "know the alphabet," not to know the letter "a." We aim then at classes with related behavior; the model is not one of appraising the "ability to jump from standing"--a domain of jumps, forward and back, sideways, landing on one foot, and so forth.... Hively and his associates have thus improved upon "classical" criterion-referencing. They have stressed the complexity of domains by pointing out subdomains....(p. 39)

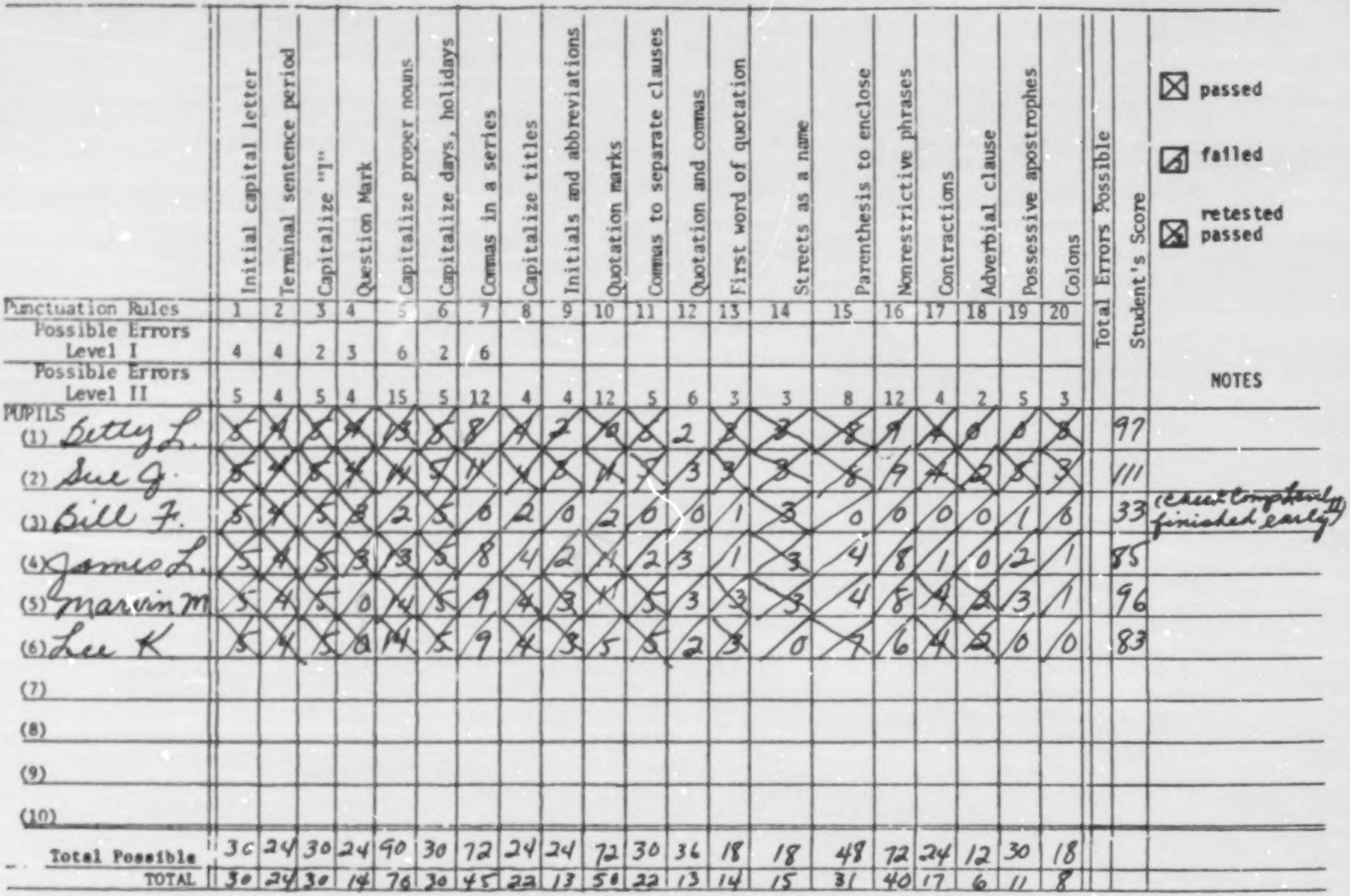
Domain-referenced testing is one approach to criterion-referenced testing in which considerable care is taken with test item selection to ensure that we can generalize from performance on the test items to the specific curriculum area (domains) that the test items are selected to represent.

Criterion-referenced testing can ready its full potential only when it is so integrated into the day-by-day functioning of the classroom that it cannot be separated out as a "testing" activity. Indeed, its contribution to the direction and programming of instructional activities should be such that the teacher sees it as indispensable for facilitating effective instruction.

The diagnostic class profile is a basic classroom management companion of a criterion-referenced test. Unlike many "clinical" tests, a school-oriented criterion-referenced test has to take into account the management problems faced by the teacher who is responsible for the total class. The basic parts of a diagnostic class profile are (a) a listing of the pupils on one axis of a matrix, and (b) a listing of the specific skill being tested on the other axis of the matrix.

An example of a diagnostic class profile used with a criterion-referenced punctuation test (Hofmeister, 1972) is listed in Figure 1. Along the top of the profile are 20 skills in a suggested order of instruction from left to right. A listing of the pupils whose skills are being evaluated is on the left-hand side of the profile. The top right-hand corner of the profile lists the coding procedure used to record test data. Most class profiles can be adapted to monitor progress as well as providing pre- and posttest data.

A well-structured criterion-referenced test should generate direct individual prescriptions for each child. In the case of the punctuation test (Figure 1), the individual prescriptions are obtained by reading from left to right after the child's name. Small group prescriptions can also be obtained by working vertically down the class profile. Figure 2 provides another example of a diagnostic class profile (adapted from Morgan, 1981). In this example the results of an instrument for assessing preacademic skills are presented in matrix form.



Punctuation Rules	Initial capital letter	Terminal sentence period	Capitalize "I"	Question Mark	Capitalize proper nouns	Capitalize days, holidays	Commas in a series	Capitalize titles	Initials and abbreviations	Quotation marks	Commas to separate clauses	Quotation and commas	First word of quotation	Streets as a name	Parenthesis to enclose	Nonrestrictive phrases	Contractions	Adverbial clause	Possessive apostrophes	Colons	Total Errors Possible	Student's Score	NOTES
Punctuation Rules	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20			
Possible Errors Level I	4	4	2	3	6	2	6																
Possible Errors Level II	5	4	5	4	15	5	12	4	4	12	5	6	3	3	8	12	4	2	5	3			
PUPILS																							
(1) Betty L.	5	4	5	4	13	5	8	4	2	10	5	2	3	3	8	9	4	0	0	3		97	
(2) Sue G.	5	4	5	4	14	5	11	4	3	11	5	3	3	3	8	9	4	2	5	3		111	
(3) Bill F.	5	4	5	3	2	5	0	2	0	2	0	0	1	3	0	0	0	0	1	0		33	(Check Comp. Level II finished early)
(4) James L.	5	4	5	3	13	5	8	4	2	11	2	3	1	3	4	8	1	0	2	1		85	
(5) Marvin M	5	4	5	0	14	5	9	4	3	11	5	3	3	3	4	8	4	2	3	1		96	
(6) Lee K	5	4	5	0	14	5	9	4	3	5	5	2	3	0	7	6	4	2	0	0		83	
(7)																							
(8)																							
(9)																							
(10)																							
Total Possible	36	24	30	24	90	30	72	24	24	72	30	36	18	18	48	72	24	12	30	18			
TOTAL	30	24	30	14	76	30	45	22	13	50	22	13	14	15	31	40	17	6	11	8			

☒ passed

◪ failed

⊠ retested passed

Class Resource Room School James High Teacher Jill Jackson Date 11/11/73

Figure 1. Diagnostic Class Profile Punctuation Test

	Billy	Jane	Sally	Melissa	Andy	Bobby	Jeannie	Tommy	Mary	Danny	Carol	Richard	Janet	Phillip	Susan
Sorts Colors	+	–	+	–	+	+	+	+	–	–	+	+	+	–	+
Identifies Colors (Names)	+	–	+	–	+	+	+	+	–	–	+	+	+	–	+
Cuts with Scissors on line	+	–	+	–	+	–	–	+	–	–	+	+	+	–	+
Colors within Boundary	+	–	+	–	+	–	–	+	+	–	–	+	+	–	–
Sorts Shapes	–	+	+	–	+	–	+	+	+	–	–	+	+	+	–
Identifies (names) Shapes	+	+	+	–	+	–	+	+	+	–	+	+	+	+	–
Draws Shapes from Dictation	–	+	+	–	+	–	–	+	+	–	–	+	+	+	–
Traces Letters	+	–	+	–	+	–	+	+	+	–	+	+	+	–	–
Copies Letters	+	–	+	–	+	–	+	+	+	–	+	+	+	–	–
Says Letters Names - Capitals	–	+	+	–	+	–	+	+	+	–	–	+	+	–	–
Writes Letters - Capitals	–	–	+	–	+	–	–	+	–	–	–	+	+	–	–
Says Letter Names - Small	–	+	+	–	+	–	–	+	+	–	–	+	+	–	–
Writes Letters - Small	–	–	+	–	+	–	–	+	–	–	–	+	+	–	–
Reads Numerals 0 - 9	+	+	+	–	+	–	+	+	+	–	–	+	+	+	–
Counts Orally 0-15	+	–	+	–	+	–	–	–	+	+	–	+	+	+	+
Writes Numerals 0-9(memory)	+	–	+	–	+	+	–	–	+	–	–	+	+	+	–
Labels (written) sets 0-9	–	–	+	–	+	+	–	–	–	–	–	+	–	–	–

A GROUP MATRIX FOR MONITORING ACQUISITION OF PRE-ACADEMIC SKILLS

Adapted from *What's going on here?: A collection of exceptional practices in Idaho.* Eugene OR: Northwest Regional Resource Center, 1977.

FIGURE 2 Sample Charts for Recording Student Performance

BEST COPY AVAILABLE

Adapting Test Content

The content of many tests is often accepted as unalterable because of association with normative testing in which alteration of either the test content or the administration procedures results in invalidating any normative information collected from such administrations. In some criterion-referenced test situations, alterations can be made without invalidating the test. In the math combinations test (Figure 3), a number of the items require negative numbers as answers. If a teacher feels that these negative number problems are inappropriate for her needs, she can delete such problems to make the test more responsive to her needs.

Random and Linear Access

In the math combinations test (Figure 3), there is very little inherent sequence among the subtraction examples. Because of that fact and the relatively low conceptual level of the task, it makes little difference where one begins teaching. This test is an example of a random access test.

There are other test areas where sequence is important, as with the subskills of the long division algorithm, or a spelling list based on word difficulty. Tests in which sequence is important are termed "linear access tests." The designation of a test as a linear or random access test has a number of instructional and testing implications. In a random access testing situation, flash cards can often be used both for testing and teaching. The use of a set of flash cards for each child is one of the quickest and most practical ways to identify and group the items a child knows and does not know. In a linear testing situation, it is not always necessary to test all the skills in the test area. For example, in spelling it might be sufficient to test until the child starts failing and then start teaching. Further testing after the child

DIAGNOSTIC ARITHMETIC

Subtraction

Date: ________ Grade: ________________ Time: ____________ Name: ________________

6	13	6	0	5	1	15	7	5	14	13
-4	-6	-2	-6	-1	-0	-8	-3	-5	-7	-9
15	5	3	6	4	11	3	11	12	10	13
-6	-2	-2	-1	-0	-8	-3	-5	-7	-9	-4
4	6	4	9	13	9	14	7	14	8	14
-2	-8	-1	-0	-8	-3	-5	-7	-9	-4	-6
0	8	2	12	8	12	16	9	7	9	10
-2	-1	-0	-8	-3	-5	-7	-9	-4	-6	-2
2	8	16	6	13	8	11	5	8	9	2
-1	-0	-8	-3	-5	-7	-9	-4	-6	-2	-11
6	17	8	7	11	16	10	11	8	3	10
-0	-8	-3	-5	-7	-9	-4	-6	-2	-9	-1
8	5	8	15	12	11	7	3	0	7	5
-8	-3	-5	-7	-9	-4	-6	-2	-4	-1	-0
10	6	10	18	12	6	11	2	9	0	14
-3	-5	-7	-9	-4	-6	-2	-6	-1	-0	-8
9	13	17	9	12	2	0	3	7	10	4
-5	-7	-9	-4	-6	-2	-6	-1	-0	-8	-3
9	15	4	10	7	8	1	3	9	12	10
-7	-9	-4	-6	-2	-11	-1	-0	-8	-3	-5

Figure 3: Math combinations test in subtraction with diagonal sequences

has started failing would certainly be aversive to the child, and very likely a waste of instructional time.

Cross-Referencing

Prescriptions based on the results of criterion-referenced tests, where possible, should cross-reference from skills to instructional materials. Skills might be cross-referenced as shown in the following example. In the example, the page numbers refer to the text being used in the child's classroom.

SKILLS	INSTRUCTIONAL EXAMPLES PAGE NUMBER	TEST EXAMPLES PAGE NUMBER
1. parts of a mixed number	172-173	174, 192, 456
2. mixed number products	176	184, 192, 456

Commercial Criterion-Referenced Testing Systems

A number of companies publish comprehensive testing systems which provide data on a child's performance on several hundred skills. Some of the systems require computer scoring services. Some of these testing systems can be cross-referenced with most of the major basal reading programs and math texts in use in the schools. Examples of some of the available criterion-referenced testing systems are: The Prescriptive Mathematics Inventory, published by the California Test Bureau, A Division of McGraw-Hill Book Co.; The Fountain Valley Teacher Support System, published by Zweig Associates; The Wisconsin Design Program, published by NCS Interpretive Scoring System; and the Beginning Assessment Test for Reading, published by Lippincott.

The following description from the preface to the Teacher's Handbook for Diagnosis: An Instructional Aide - Mathematics, Levels A and B, (developed and published by Science Research Associates, 1980) is representative of the approach taken in some of the well-developed commercially available criterion-referenced testing systems:

> Each diagnosis lab provides sets of diagnostic probes-- exercises designed to help the teacher identify the specific weaknesses of a student. The exercises are based on the comprehensive lists of learning objectives on the back of the probe. By freeing the teacher from the time-consuming task of developing detailed diagnostic tests, a diagnosis lab enables him to focus his efforts on the instruction or remediation needed by each student to correct identified weaknesses.
>
> Also provided is a Prescription Guide with which teachers and students can quickly determine where--in a large number of texts and supplementary learning kits--materials pertinent to the achievement of each objective are presented. Multiple references are given for every learning objective. Thus, if the student has had difficulty with a given text presentation or other instructional material, he/she can be directed to another. This approach allows for a multitext and multimedia approach, accommodating a wide variety of learning and teaching styles.
>
> While some educators still champion "tacit learning" and stress the impossibility of reducing all aspects of the educational process to behavioral objectives and demonstrable outcomes, it is generally accepted that mastery of fundamental skills is, perhaps, best taught in this fashion. This is the basic premise of the series of diagnostic labs; to provide a convenient method of helping students attain all the fundamental skills and concepts upon which much of their future education will rest. (p. 1)

Setting a Criterion

One of the major purposes of a criterion-referenced test is to help determine which skills to incorporate in remedial instruction. The manual accompanying a criterion-referenced test should specify skills to be tested and the sequence in which these skills should be taught. As yet, not enough research has been done to indicate precisely what degree of success should be used to indicate when a skill is mastered. The following are guidelines for setting a criterion:

1. A criterion should be based on the subject matter, and one level of attainment (e.g., 95 percent correct) should not be expected to apply to all subject matter areas.

2. The range used should be 80 percent to 100 percent. If a skill is important in terms of personal safety (for instance, driver's training) or as a prerequisite to an important skill, a range of 90 percent to 100 percent would be appropriate.

3. The criterion should be related to the nature of the test. If the test item was designated as follows:

Question (1) 34
-19

then a criterion of 80 percent on the total test might be appropriate. If the test question were of the form:

Question (1) 34 Check one: (a) 13
-19 (b) 14
(c) 15
(d) 16

then a criterion of 90 percent might be more appropriate because of the chance factor and prompting associated with the multiple choice format.

4. A criterion may take into consideration the performance of others. A level may be established by referring to the relative position the student holds in a particular group. This adds a norm-referenced element to the test.

5. A level may be set by judging minimal competence. Experts decide what score a minimally competent person should obtain.

Use of Profiles in Planning Educational Programs

Teachers and other school personnel are often interested in the relative levels of mastery that a student demonstrates in various skill areas and/or in the differences in performance across skill domains within one particular test. Although the construction of profiles from test scores is a common practice, their analysis and interpretation are complex procedures. School personnel must be cautious in using such information for purposes of educational diagnosis and program planning.

Salvia and Ysseldyke (1978) note that "educators expect children with above-average intelligence to perform better than average in their academic work. If their achievement is not also above average, it is often a source of concern" (p. 408). It is also true that children who achieve relatively low scores on intelligence tests are expected to perform below average. Too often, poor performance on a test of ability leads to expectations of poor performance in other skill areas.

Professionals have used assumptions about the correlation between ability and achievement to support classification of students in different areas of exceptionality. Salvia and Ysseldyke (1978) point out that:

> Flat profiles of individuals whose functioning measures significantly below average in both intelligence and adaptive behavior are used to confirm diagnoses of mental retardation....A child who has a significant discrepancy between measured intelligence and both measured achievement and perceptual or language functioning, or both, may have a learning disability. (p. 409)

It is also assumed that program planning may be facilitated through use of profiles of achievement tests. Poor performance in one or more academic areas may indicate a need for additional instruction. Conversely, high achievement in an academic area may indicate special interests or skills which can be capitalized on in instruction.

If profile analysis is to be of use in developing instruction, attention must be given to the reliability of score differences and to the differences in norm samples. Salvia and Ysseldyke (1978) provide an example of profile analysis which is reproduced in Appendix C.

Activity

Select a norm-referenced achievement test commonly used in schools:

1. Analyze the contents for adequacy in assessing a specific ability of students (e.g., multiplication of two-place numerals, inferential comprehension of grade level reading material).

2. Compare the test content with the scope and sequence of the curriculum being used in the school.

3. Identify specific changes in instrument content and test procedures (e.g., use of oral responses, written responses) that allow for more precise recommendations to be expressed in terms of the curriculum and materials being used in the pupil's school.

Evaluation of Intervention Effects

The success of any effort to improve a student's academic and/or social performance should be evaluated from two aspects: (1) The extent to which the treatment results in significant change and/or progress, and (2) the extent to which the treatment program is valid.

Evaluation of Intervention Outcomes

Public Law 94-142 requires both assessment of students who are receiving special education service and educational interventions and the evaluation of the effectiveness of those services in achieving objectives for the students. The current emphasis on individual educational plans (IEPs) has resulted in a need for the school personnel to become involved with assessment of outcomes of their programs for handicapped students. Beyond this mandated assessment of the programs of a relatively few of the students in regular classrooms, much benefit may be gained from expanding program evaluation to assess the progress of all students.

In his discussion of the assessment of behavioral change, Phye (1979), points out that the educational professional is concerned both with the nature of the change and with the evaluation of the impact of the intervention.

Nature of the change. Here, the types of questions to be asked deal with the reliability of the data showing a change in behavior and the direction of the change. Evaluation in this area is typically concerned with the improvement and/or modification of the intervention program. A standard method of assessing the nature of a change due to intervention is through a simple pretest-posttest procedure. In addition, the assessment of instructional objectives must be dealt with on a continuing basis at all levels of an instructional task in order to monitor instructional methods and materials.

Evaluation of intervention impact. To evaluate the effectiveness of the educational program the following question should be asked: "To what extent was the observed change a result of the intervention treatment?" Here the program to be evaluated may be any program or curriculum by which a student's education is guided. The "treatment" may refer to the standard program in a regular classroom or a modified program in which an individualized educational plan is carried out.

Given the ethical and practical restrictions that are placed on the professional in the school setting, evaluation of program effectiveness via "true" experimental design is not generally possible. Phye (1979) therefore suggests that some form of time-series design would be valuable in evaluating program impact. In other words, an on-going evaluation of a program may not be able to demonstrate that a program is necessarily better than other possible programs, but it should be able to demonstrate that students in that program are progressing over time.

Treatment validity. Closely related to the issue of evaluating program effectiveness is that of treatment validity. It is common practice in special education to use certain tests as a basis for prescribing educational treatment. When tests are used in this way, it is necessary to assess their treatment validity, i.e., to assess the degree to which they enable educators to accurately describe a program's effectiveness. In this section a procedure is described for assessing the treatment validity of a test.

There are essentially three areas of possible threat to treatment validity. These areas were described by Gallery and Hofmeister (1978) as: (a) the lack of a relationship between test and treatment; (b) the low quality of treatment, and (c) the lack of relationship between treatment and curriculum.

Relationship between test and treatment. Can an educational treatment be prescribed, given the test results? If the answer is no, decisions about treatment are unreliable. The treatment decisions may result in the student's failure to achieve, and such failure would be difficult to assess since it could be the result of inaccurate measurement, ineffective treatment, or both. Related to this concern is the teacher's ability to determine a starting point for treatment. If the test results do not give some indication of where the treatment should begin, the treatment may be inefficient or ineffective.

Quality of treatment. Was the educational program field-tested or researched? There must be empirical evidence for an assumption that the program should yield affirmative answers to the following questions:

1. Is there evidence that the skills being targeted in the program were mastered by students on whom the materials were field-tested?

2. Are the responses required in the student assessment similar to those required in the educational program?

3. Are the students on whom a program was field-tested similar to the pupil(s) being assessed?

Relationship between treatment and curriculum. To what extent are the program objectives contained in the curriculum? The issue here is one of efficient use of time. A discrepancy between program objectives and the curriculum content could mean that the program has little relevance to the student's needs.

Validity involves all interpretations of the assessment data. Validity does not necessarily refer to the nature of the test(s) being used. That is, one cannot automatically assume treatment validity for a criterion-referenced test and for a norm-referenced test. The extent to which test results can assist a professional concerned with planning effective educational programs depends on the degree to which the criteria outlined are met.

Conclusion

This overview of educational assessment and its purposes has been designed to assist the progress of the diagnostician through the complex maze of educational decision-making.

Educational assessment is a multidimensional procedure by which the infinite variety of a child's behaviors is observed and evaluated. Legal and social mandates, combined with research findings, delineate the responsibilities of school professionals in assessment. These requirements have resulted in a crucial need for educational decision-makers to understand the nature and limitations of assessment technology.

The goal of educational assessment is the generation of educationally relevant decisions designed to make positive differences in the lives of children. School personnel must be prepared to be accountable for such decisions. One measure of such accountability will be the valid and reliable relationship between assessment, the demonstrated needs of the student, and the success of instructional strategies designed to provide for those needs. To assess the pupil and ignore the instructional environment is to imply that the pupil is totally responsible for his/her level of educational attainment.

In this module assessment strategies and issues have been covered in a very cursory manner. Obviously, adequate competence in pupil assessment requires much more than familiarity with the contents of this module. Considerable practice and experience is required to develop assessment skills. Some of this practice and experience should be a part of all classes that deal with methods of instruction, whether these deal with reading methods for elementary school teachers or science methods for high school teachers. Some of this experience should be gained in well-selected practicum placements.

Attached in the appendices of this module are a few exercises which could be used in teacher education programs. Of course they are only a beginning. What is crucial is that teachers leaving today's colleges have a background in and commitment to on-going student evaluation. It is only then that we can feel confident that, in an era of growing heterogeneity in school populations, an appropriate education will be provided for all of America's children and youth.

APPENDIX A

Posttest

Posttest

1. According to Public Law 94-142, the primary purpose of assessment is to

    a. develop appropriate educational programs for handicapped children.

    b. make educational decisions about the placement of children in curricula and programs.

    c. evaluate the effectiveness of the child's individual education program in meeting identified goals and objectives.

    d. all of the above.

2. Because school professionals have different training and unique skills, the responsibility for educational assessment should be

    a. allocated to the person who spends the most time with the child in educational settings.

    b. the sole responsibility of the school psychologist as a function of his/her training.

    c. an integrative process involving every person who has contact with the child in the educational setting.

    d. the co-equal responsibility of those who are in a position to contribute educationally relevant information about the child.

3. Continual evaluation of the child's progress is necessary for the following reason(s):

    a. to make sure the assessment instruments used reliably identified the general nature of the child's problems.

    b. to ensure that the child is not causing further problems for the teacher.

    c. to ensure that the student is in the right curriculum and is learning at his/her maximum rate.

    d. to ensure that the student's placement will be final and that there will be no threat of his/her personal adjustment because of movement to a new position.

4. In using the information from observation and interviews, school personnel

   a. can view the information as accurate because of the reliability and validity of the instruments used.

   b. know that they have obtained accurate information.

   c. know that the information they obtained is more accurate than information from only the child's teacher because the latter is not specially trained to assess students.

   d. know that interview-observation techniques are too subjective to yield precise and reliable information, but can be used to supplement and verify other data gathered on the child.

5. A criterion-referenced test places a child in terms of

   a. a sequence of tasks.

   b. a set of percentiles.

   c. other children.

   d. a set of norms.

6. Norm-referenced tests are often

   a. sensitive to a specific curriculum.

   b. good diagnostic instruments.

   c. insensitive to a specific curriculum.

   d. useful for directing specific remedial programs.

7. A domain-referenced test is one form of a

   a. standardized diagnostic test.

   b. norm-referenced test.

   c. standardized achievement test.

   d. criterion-referenced test.

8. The intelligence test is a

   a. norm-referenced test.

   b. criterion-referenced test.

   c. domain-referenced test.

   d. achievement test.

9. If a criterion-referenced test was used to rank the pupils in a class, the test would be used as

   a. a criterion-referenced test.

   b. an achievement test.

   c. a norm-referenced test.

   d. a domain-referenced test.

10. There is a strong movement to de-emphasize testing for purposes of

   a. identifying and classifying.

   b. remediation.

   c. specifying students' present level of skills.

   d. planning educational programs.

11. The most used special education testing procedures are those procedures which

   a. give a percentile ranking.

   b. facilitate remediation.

   c. facilitate giving end-of-term grades.

   d. place a child in relation to his peers.

12. When a child meets criteria on a domain-referenced test, he should

   a. have mastered the test items only.

   b. be able to master all examples of the domain when encountered.

   c. have mastered at least half of the possible examples of that domain.

   d. none of the above.

13. When selecting an assessment instrument, professionals should first

a. decide on the scoring procedure.

b. make sure that there are at least two forms of the test.

c. clarify the purpose of the test.

d. make sure the students can understand the directions.

14. For criterion-referenced testing programs to be effective,

a. they should be administered by an impartial outsider.

b. they should be administered at the end of the year.

c. they should be integrated into the daily teaching procedures.

d. they should be administered at the beginning of the year.

15. A matrix with one axis listing the skills and the other listing the pupils is a

a. table of norms.

b. diagnostic class profile.

c. individual prescription.

d. standardized test.

16. A criterion-referenced test can be made more practical by

a. adding norms.

b. expressing results as a percentage.

c. cross-referencing the skills.

d. expressing results as ranks.

17. A diagnostic class profile can be used for

a. recording pretest data only.

b. recording pre- and posttest data and monitoring student progress.

c. recording posttest data only.

d. recording the data on which each class is given.

18. A well-constructed criterion-referenced test should

a. generate prescriptions for direct individual instruction.

b. generate percentile rankings.

c. provide information regarding the appropriate norm-referenced test to give.

d. none of the above.

19. If a teacher feels that parts of a criterion-referenced test are inappropriate, he/she should

a. not use that particular test.

b. adapt the test by deleting the parts he/she feels are inappropriate.

c. make up his/her own test.

d. delete those parts he/she feels are inappropriate, but realize that the test is no longer valid.

20. Test developers claim that Test A is more reliable than Test B. To determine the accuracy of their claim, what information would you need about Test A?

21. The authors of the WRAT state that no attempt was made to obtain a representative national sample of students for the standardization of the test. Each level of the test was standardized on at least 150 males and 150 females at each of nineteen age levels, producing a total standardization population of 5,868 persons at Level I and 5,933 persons at Level II. Norms were not stratified on the basis of race, ethnic group membership, socio-economic level or geographic region. Schools in only seven states were included in the standardization sample. No handicapped children were included (Salvia & Ysseldyke, 1978, p. 160).

Identify five characteristics students may possess that require caution in interpreting test results and that may demand adaptation in test administration.

22. What are the two questions to ask when evaluating the nature of behavioral or academic change in a student during and after intervention?

23. When assessing treatment validity, which three areas will be of possible threats to validity?

24. How do the needs of administrators in regard to assessment information differ from the needs of teachers? What are the practical implications of these differing needs?

APPENDIX B

Exercises

Exercise 1

1. It has been noted in the research literature that the lack of mastery of sounds of letters is a major skill deficit in many poor readers. You will be presented with a paragraph which a student has read with markedly poor skill. Specific problems are identified in the test by a code which appears beneath the paragraph. Look for a consistent pattern and frequency of errors. Design a criterion-referenced test which will cover all the problems identified. The test will be individually and orally administered as a pre- and posttest to monitor instruction in the pronunciation deficits.

2. Also prepare a diagnostic class profile to be used in association with the test. The class profile should include the following components:
    a. a space for listing at least 10 pupils;
    b. a listing of the sound being evaluated, and
    c. a clear and simple coding procedure for entering the test results. The coding system should allow for updating the profile as skills are mastered.

   All the administrative procedures should be described in sufficient detail so that another teacher might use the instrument without additional information and/or training.

Reading Paragraph

VAN'S CAVE

(very)
(o) ∧ c (pronounced "d") c (o) (o) (big)
Van is an old, old man. He lives in a cave near a lake. Van has a ∧

(o) (said "bed" at first) c
bed and a stove in the cave. Van made the bed of old lumber. He made the stove

(said "bed" at first) c
of old bricks. He sleeps in the bed. Spot likes to sleep near the stove. His

bed is just an old coat.

(a) (līve) (o)
Spot likes to live in Van's cave. It is home for the old man and the dog.

(o) (a) (o) (a)
Once Van and Spot went to hunt ducks. Van saw a flock of ducks land on the lake.

(fird) (fird) (o)
He fired and fired at the ducks. "I hit five ducks," said Van. "Jump into the

(a) (a)
water, Spot, and get the ducks for me." Spot jumped into the waves and swam to

get the ducks.

Scoring Code

1. ―― (letter or word sounded incorrectly)
2. ○ (word skipped)
3. c (word or letter which student self-corrects)
4. ∧ (insertion)
5. sk (rows skipped)

Exercise 2

You will be presented with a case report of a student who has been referred to a school psychologist for evaluation. On completion of the reading, decide whether you have sufficient information about the child to make recommendations for him as regards educational placement and treatment. If you do not feel that there is sufficient information, what additional data do you require and how will you obtain it? Detail the procedures you would follow and the methodological considerations you would make in your choice of assessment instruments. Include an outline of the procedures you would follow in evaluating the success of any educational program you might recommend. Discuss the result of this exercise with other members of the group.

Case Report

Client: Wadell P.

Age: 10.7

Birthdate: 2/3/65

Date of Evaluation: 3/10/75

Parents: Edward and Jane P.
327 East 800 North Boulevard
Los Angeles, California

Referral Reason:

Wadell P. is a black student who was referred by his teacher, Mrs. Sharon K., and Principal, Mrs. Noreen S., of A________ School, for psychological evaluation and possible placement in a service plan for an educationally significant handicap.

Tests Administered: Stanford-Binet, LM, Intelligence Test
Stanford Achievement Test
Draw-a-Man (unscored)

Background Information:

Wadell P. transferred to A______ School on January 7, from a school in Los Angeles, California. His recores were not available from that school at this time, but his current principal, Mrs. S____, related that Wadell's mother indicated he "had some problems" at the previous school. He was placed in the classroom of Mrs. K______, and over the past month she has voiced some concerns which necessitated this referral. Specifically, Mrs. K______ reported that Wadell does not seem to be interested in making friends or in the activities within the classroom. She describes his behavior as "impulsive" but generally "underactive, as he seems to act in slow motion." Often he "refuses to talk," and often fails to make appropriate responses in social situations (i.e., doesn't talk to other children when spoken to, doesn't join in games). She also indicated that "he is a clumsy child, or at least uncoordinated for his age. When you can get him to respond, he seems to be a little behind in all subjects, but especially reading."

On the first day of the evaluation, Mrs. P_____ was in attendance and provided the following information. Wadell is the second of four children. He has had no significant illnesses up to this time. When tested previously, his vision was believed to be 20/20 in both eyes and his hearing was "OK." Mrs. P______ stated, "I don't think there is a problem with Wadell; he's fine at home." Then was added, "Mr. P______ is employed as a plumber, presently, and the family receives assistance through the Department of Family Services." Both parents reportedly completed the ninth grade. Mrs. P______ worked as a sales clerk "before getting married and becoming a housewife." Mrs. P_____ reported that both parents take little interest in community affairs.

Behavioral Observations:

The tests were administered in two morning sessions on two consecutive days. On the first day, Wadell was brought to the testing room in the school by his teacher. When introduced to the examiner, Wadell remained quiet. At this time, Wadell was only given a Draw-a-Man, which was not scored, and whose purpose was only as a rapport builder. On the second day of testing, Wadell was frequently verbal with the examiner, and his speech could be characterized as "slow." During administration of the Stanford-Binet (an I.Q. test), some of the answers were impulsive, especially for the memory-for-digits items. Although rapport could be characterized as good, Wadell seemed quite distractable from the task at hand when the examiner would reach into the kit for the next item. Wadell remained in his seat except on two occasions. When walking, he seems to drag his feet in an uncharacteristic gait. He does not hold a pencil solidly.

Test Results:

The Stanford Achievement Test was administered first as a broad-based assessment of Wadell's academic achievement up to this point. He was administered the Primary I Battery (or Intermediate II Battery) and according to the norms supplied, obtained the following scores:

a. Word Reading: one year behind	a. same
b. Paragraph Meaning: 10 months behind	b. same
c. Word Study Skills: 10 months behind	c. same
d. Vocabulary: 4 months behind	d. same
e. Spelling: 5 months behind	e. same
f. Arithmetic: 6 months behind	f. Arithmetic (1) Concept Formation 3 months behind (2) Computation 3 months behind (3) Application 5 months behind

The Stanford-Binet was administered next as an indicator of Wadell's abstract verbal learning and problem-solving ability. The test results were considered to be valid as Wadell was passively cooperative. According to the norms supplied, Wadell received an IQ=77. He had particular problems on the vocabulary, maze, and memory for digits items. He was unable to give a reasonable answer when asked, "What would you do if you were asked by a stranger how to find the nearest service station?" (Answer: "Who? What service station?"). Scattered throughout the other items were answers which can only be characterized as a puzzled look, as if Wadell just didn't understand the question. It is felt that Wadell has difficulty with receptive language. Questions had to be repeated; he did not engage in conversation.

Summary

Wadell's assessment indicated that he is performing at a low level academically, particularly in receptive and expressive oral and written language. It should be noted that when a social worker visited Mrs. P_______ at home, she found that Wadell's behavior at home is in sharp variance with his behavior at school. He is capable, happy, and talkative with his peers. No one in the family is concerned about his pattern of poor performance at school.

Exercise 3

Take a commercial elementary school text in arithmetic.

1. Develop a criterion-referenced test that covers the major skills which students are expected to learn while working in this text.
2. Design a record that would allow a teacher to keep a sequenced, on-going record of student progress while working in this text (indicating your criterion for mastery of specific skills).
3. Develop a brief screening test which would allow you to
   (a) determine readiness of a student to use the text, and
   (b) to place students at the appropriate place within the text.

Exercise 4

It has been noted in the research literature that the lack of mastery of the sounds of letters is a major skill deficit in many poor readers. Design a criterion-referenced test that would be individually and orally administered. The test should cover:

(a) all the consonants,

(b) all the short vowels,

(c) all the long vowels, and

(d) several common consonant blends.

Also prepare a diagnostic class profile to be used in association with the test. The class profile should include the following components:

(a) a space for listing at least 10 pupils,

(b) a listing of the sound being evaluated, and

(c) a clear and simple coding system should allow for updating the profile as skills are mastered.

All the administrative procedures should be described in sufficient detail so that another teacher might use the instrument without additional information and/or training.

Exercise 5

The situation: A high school remedial math teacher set up a grading system for the end-of-term reports as follows: A = 90%, B = 80%, C = 70%, D = 60%. The percentages refer to the average of the mid-term and final exam results. How might the teacher apply some of the mastery learning constructs and still award grades?

Responses to Exercises

1. You will note that there is a significant pattern (or set of patterns) to errors made in reading this paragraph. There are numerous cases of the reversal of the letters "o" and "a" in different words. Some words were substituted for those in the test, and there was a constant omission of "the" and "an," though significantly not of "a." Words containing "i" were mispronounced. Errors in the second paragraph indicate that the student did not discriminate the word, "stove." He/she seemed to expect that every reference would be to "bed." However, he/she did make a correction on both occasions. Any remediation should focus on these deficits. Your criterion-referenced test should be constructed to monitor progress in these areas. When constructing the test, remember to include criteria for mastery and to rationalize your choice.

2. Information presented in this case study is not adequate for you to make an appropriate decision about educational placement and treatment for Wadell.
    a. There seems to be a discrepancy between Wadell's behavior at home and at school. You would need more information about such variables as the teacher's expectations of Wadell, demands made on him by the classroom environment, and the nature of reinforcement methods employed in the classroom.
    b. The testing instruments used should be evaluated to determine if they were appropriate, particularly in terms of language, concepts tested, relationship to curriculum, etc.
    c. You would recommend that further assessment be conducted. A thorough assessment of language, both receptive and expressive, is necessary.
    d. You would recommend further assessment in deficit academic areas, with particular focus on criterion-referenced information. You would, of

course, outline the considerations you would make in selecting the assessment instruments.

e. Following the assessment process, you would examine the curriculum in which Wadell was placed to determine which deficit areas are the results of inadequate or inappropriate classroom learning experiences and which are the results of learning problems.

3. Responses will vary according to the text selected. Students may benefit from comparing responses with students working from the same text. Comparisons of criterion-referenced tests based on different texts designed for the same grade are often useful in showing how curriculum can affect a student's preparedness for standardized, norm-referenced achievement tests.

4. The attached list of basic sounds might be used as a resource for this exercise if the participants do not have curriculum knowledge in this area. The major point to be brought out here would be the alternative of assigning grades in terms of units mastered rather than in terms of average scores. For example, a grade of "A" might be assigned to students who master nine of the ten educational units in a given course.

Supplement for Exercise 4

(Single Letters)

Sound Number	Sound	Key Words
1	m	me, him
2	a	apple, fat
3	s	so, kiss
4	* ē	mē, hē
5	f	full, if
6	d	did, feed
7	r	ran, rip
8	i	it, pin
9	c	cat, traffic
10	o	ox, not
11	n	not, sun
12	t	top, hit
13	h	hat, how
14	u	up, rug
15	g	go, rag
16	l	lit, will
17	w	win, watch
18	k	kiss, milk
19	* ō	ōpen, gō
20	v	very, live
21	p	pat, sip
22	e	end, pet
23	b	big, rib
24	y	yes, yellow
25	j	jump, jeep
26	x	box, six
27	* ȳ	mȳ, crȳ
28	z	zoo, buzz

Letter Combinations

Sound Number	Letter Combination	Key Words
29	ir	bird, fir
30	th	then, that
31	qu	quit, queen
32	er	her, faster
33	ea	seat, mean
34	ol	old, bolt
35	ur	turn, urn
36	ee	feet, need
37	wh	when, why
38	oo	boot, moon
39	aw	jaw, flaw
40	or	for, order
41	ai	maid, pain
42	oa	oat, load
43	ou	out, loud
44	igh	night, sigh
45	ay	stay, pay
46	al	bald, fall
47	oy	boy, toy
48	sh	shout, wish
49	ar	car, start
50	oi	oil, point
51	ch	chop, rich
52	** a-e	gate, plane

Letter Combinations
(Continued)

Sound Number	Letter Combination	Key Words
53	** i-e	like, dive
54	** o-e	rode, home
55	** u-e	mule, flute
56	ew	news, skew

* ē, ō, ȳ, - the line over these letters indicate that the letters be the long sounds and have the same sound as the name of the letter (e.g., as in bē, gō, crȳ, hē, nō, drȳ)

** - a-e, i-e, o-e, u-e } These are long sounds (e.g., "a-e" is the long sound of "a") followed by a "silent e" (e.g., as in make, home, drive, fuse)

APPENDIX C

Keys to Pretest and Posttest

Key to Pretest

Short Answer Items

1. To make educational decisions about the performance of students.
2. Testing is only one method of assessing a child. Assessment is a flexible, continuous process leading to an ongoing program which may be modified in the interests of the student's life situation and of a reduction in his/her current specific difficulties.
3. a. Selection and administration of tests in a nondiscriminatory manner.
    b. Provision of test materials and procedures in the child's native language or mode of communication.
    c. Prohibition of the use of a single test or procedure as the sole criterion for determining educational placement.
4. a. Child's files.
    b. Information from significant others in the child's life.
    c. Observation of the child.
    d. Norm-referenced tests.
    e. Criterion-referenced tests.
5. a. Current life circumstances.
    b. Developmental history.
    c. Extra-personal factors.
    d. Situational factors.
6. a. Referral.
    b. Screening.
    c. Evaluation of screening and planning interventions.
    d. Evaluation of intervention effects.

7. a. Observer bias.

   b. Awareness by the student that he/she is being assessed.

   c. Inconsistency between interviewers.

   d. Test invalidity and/or unreliability.

8. A student's scores on an achievement test reflect pupil ability as compared to others of comparable age level and the instructional effectiveness of an educational program.

9. a. Use of the test for purposes other than those for which it was designed.

   b. Use of the test for individual purposes when designed for use with a group.

10. a. To provide recommendations for instructional and behavioral goals.

    b. To provide recommendations for educational placement.

    c. To measure present skills.

    d. To measure progress toward goals.

    e. To measure the effect of an instructional program.

11. Norm-referenced tests evaluate the child's performance by comparison with that of other children. Criterion-referenced tests evaluate the child in terms of some learning task at some absolute standard that is independent of other children's performance.

12. A domain-referenced test is one in which the emphasis is placed on precisely identifying skill areas (domains). The test items of a domain-referenced test are selected so that one can be certain that the child who meets criteria on the items which represent a domain (such as long-division of decimal fractions using single digit divisors) would be able to meet all the other examples of a domain when they were encountered.

13. a. Criterion should be based on subject matter.

    b. The range should be 80-100 percent.

    c. The criterion should relate to the nature of the test.
    d. Performance of others.
    e. Expert decision on minimal competence.

14. Nature of the change in the student and impact of the intervention.

15. The degree to which tests enable educators to accurately describe a treatment effect.

16. a. Parental permission is necessary before initiation of diagnosis or evaluation and for placement beyond a regular education classroom.
    b. Multidisciplinary evaluation.
    c. Due process or procedural safeguards to ensure the educational rights of the child.

True/False Answers

1. F
2. F
3. T
4. F
5. F
6. F
7. T
8. T
9. F
10. F
11. T
12. F
13. T

Key to Posttest

1. d
2. d
3. a
4. d
5. a
6. c
7. d
8. d
9. c
10. c
11. b
12. b
13. c
14. c
15. b
16. c
17. b
18. a
19. b

20. a. Validity of the test for the purpose and the populations.

    b. Information about technical adequacy of test
    (1) administration
    (2) standardization/norming groups.

    c. Knowledge of the physical and mental characteristics of the child being tested.

    d. Knowledge of curriculum in which child is currently placed.

    e. Reliability coefficients (for test and subtests).

21. Answers should include: Characteristics of geographic region, cultural/ethnic background, age of student (C.A. and M.A.), nature of disability and its possible effect on nature and rate response; any medication prescribed for the child and its influence on the nature and rate of response.

22. a. What is the reliability of the data supporting a change?
    b. What is the direction of the change?

23. a. The lack of relationship between test and treatment.
    b. The quality of treatment.
    c. The lack of relationship between treatment and curriculum.

24. Administrator Needs

Student program placement.
Evaluation of educational progress.
Appraisal of effectiveness of specific curricula and programs.
Identification of students with specific disabilities for funding purposes.

Teacher's Needs

Instructionally relevant information.
Precise knowledge of students' specific skills and deficits.
Any other information that will facilitate effective intervention.

Practical Implications

Nature of assessment instruments (norm-referneced, criterion-referenced, diagnostic rather than achievement test, relationship between test and curriculum sequence).

References

Adelman, H. S., & Taylor, L. Initial psychoeducational assessment and related consultation. Learning Disability Quarterly, 1979, 2, 52-64.

Bloom, B. S. The new direction in educational research: Alterable variables. Phi Delta Kappan, 1980, February, 382-385.

Burton, N. W. Societal standards. Journal of Educational Measurement, 1978, 15(14), 263-271.

Capie, W. Sampling considerations in classroom research. Paper presented at the meeting of the American Educational Research Association, Los Angeles, April 1981.

Clarke, A. M., & Clarke, A. D. B. Mental deficiency: The changing outlook. New York: Free Press, 1975.

A Common Body of Practice for Teachers: The Challenge of Public Law 94-142 to Teacher Education. Washington, D.C.: American Association of Colleges for Teacher Education, 1980.

Diagnosis: An instructional aide - mathematics, levels A and B. Chicago: Science Research Associates, 1980.

Donlon, T. F. Referencing test scores: Introductory concepts. In W. Hively & M. C. Reynolds (Eds.), Domain-referenced testing in special education. Reston, VA: The Council for Exceptional Children, 1975.

Engelmann, S., Granzin, A., & Severson, H. Diagnosing instruction. Journal of Special Education, 1979, 13(4), 355-363.

Forness, R. F., Thornton, R. L., & Horton, A. A. Assessment of applied academic and social skills. Education and Training of the Mentally Retarded, 1981, 16(2), 104-109.

Gallery, M., & Hofmeister, A. M. A method for assessing the treatment validity of tests in special education. The Exceptional Child, 1978, 25(2), 105-113.

Glaser, R. Instruction, technology, and the measurement of learning outcomes: Some questions. American Psychologist, 1963, 18(8), 519-521.

Hickey, M. E., & Hoffman, D. H. Diagnosis and prescription in education. Educational Technology, 1978, 13(10), 35-37.

Hively, W. Introduction to domain-referenced testing. Educational Technology, 1974, 14(6), 5-10.

Hofmeister, A. M. Diagnostic capital letters and punctuation tests. Logan, UT: Exceptional Child Center, Outreach and Development Division, Utah State University, 1972.

Hofmeister, A. M. Testing and treatment validity. Journal of Learning Disabilities, 1979, 12(3), 73-75.

Hofmeister, A. M., & Espeseth, K. Predicting academic achievement with TMR's. American Journal of Mental Deficiency, 1970, 75(1), 105-107.

[illegible], J. L., & LeGear, L. An emerging model: Criterion-referenced evaluation. Thrust, April, 1973.

Kamin, L. J. Social and legal consequences of IQ tests as classification instruments: Some warning from our past. Journal of School Psychology, 1975, 13, 317-323.

Kavale, K., & Mattson, P. D. One jumped off the balance beam: Meta-analysis of perceptual motor training. Riverside: University of California at Riverside, School of Education, 1980.

Lien, A. J. Measurement and evaluation of learning. Dubuque, IA: W. C. Brown, 1971.

McIntyre, K. H. Adaptive behavior. In A. L. Pasanella, T. B. Sales, & C. B. Volkmor (Eds.), Cross-cultural educational assessment skills workshop: Training processes and materials. Los Angeles: California Regional Resource Center, 1978.

Mercer, J. Labeling the mentally retarded. Berkeley: University of California Press, 1973.

Mischel, W. Personality and assessment. New York: Wiley, 1968.

Morgan, D. I. A primer on individualized education programs for exceptional children. Reston, VA: The Foundation for Exceptional Children, 1981.

Oakland, T. (Ed.). Nonbiased Assessment (a series of training modules). Minneapolis: University of Minnesota, Department of Psychoeducational Studies, 1981.

Phye, G. School psychologists as consultants in the evaluation of learning and intervention outcomes. In G. D. Phye & D. G. Reschly (Eds.), School psychology: Perspectives and issues. New York: Academic Press, 1979.

Poteet, J. A. Informal assessment of written expression. Learning Disability Quarterly, 1980, 3(4), 88-98.

Reynolds, M. C. Trends in special education: Implications for measurement. In W. Hively & M. C. Reynolds (Eds.), Domain-referenced testing in special education. Reston, VA: The Council for Exceptional Children, 1975.

Salvia, J., & Ysseldyke, J. E. Assessment in special and remedial education. Boston: Houghton-Mifflin, 1978.

Shertzer, B., & Linden, J. D. Fundamentals of individual appraisal: Assessment techniques for counselors. Boston: Houghton-Mifflin Co., 1979.

Spivack, M., & Swift, G. *Devereaux Elementary School Behavior Scale*. Devon, PA: The Devereaux Foundation, 1967.

Thurlow, M. G., & Ysseldyke, J. E. Current assessment and decision-making practices in model LD programs. *Learning Disability Quarterly*, 1979, *2*(4), 15-24.

Wechsler, D. *Manual for the Wechsler Intelligence Scale for Children, Revised*. New York: Psychological Corporation, 1974.

Wehman, P., & Schleien, S. Classroom techniques. *Education and Training of the Mentally Retarded*, 1980, *15*(1), 50-57.

Wysocki, T., Fuqua, W., Davis, V. J., & Breuning, S. E. Effects of Thioridazine (Mellaril) on titrating delayed matching-to-sample performance of mentally retarded adults. *American Journal of Mental Deficiency*, 1981, *85*(5), 539-547.

Ysseldyke, J. E., & Algozzine, B. Perspectives on assessment of learning disabled students. *Learning Disability Quarterly*, 1979, *2*(4), 3-13.

Recommended Reading

Adelman, H. S., & Taylor, L. Initial psychoeducational assessment and related consultation. *Learning Disability Quarterly*, 1979, *2*, 52-64.

Abstract: Criticism of current approaches to assessment indicates the need for effective alternatives. Given a general understanding of the criticism, the authors describe the basis for an alternative approach to initial assessment activity. Specifically, the discussion focuses on (1) a conceptualization of initial assessment and consultation; (2) the problem-solving paradigm as a framework for guiding this activity; and (3) a description of procedures and initial data from a demonstration program.

Becker, W., & Engelmann, S. *Teaching 3: Evaluation of Instruction*. Chicago: Science Research Associates, 1976.

Abstract: This book elaborates and exemplifies many of the constructs presented in this section on educational assessment. Included in this book are the following topics: kinds of tests; understanding test scores; constructing instructional-program-based tests; approaches to monitoring student progress; outcome evaluation with criterion-referenced tests; norm-referenced achievement tests; interpreting norm-referenced test scores; and who can be taught.

Gallery, M., & Hofmeister, A. M. A method for assessing the treatment validity of tests in special education. *The Exceptional Child*, 1978, *25*(2), 105-113;

Hofmeister, A. M. Testing and treatment validity. *Journal of Learning Disabilities*, 1979, *12*(3), 73-75.

Abstract: The above two articles include specific recommendations to the psychologist who is concerned about selecting tests that will have instructional relevance. A detailed frame of reference listing specific criteria for test selection is provided.

Learning Disability Quarterly, 1979, *2*(4).

Abstract: This whole issue highlights the assessment of learning disabilities. Three articles that are especially applicable have been included in the Annotated Bibliography.

Salvia, J., & Ysseldyke, J. E. *Assessment in special and remedial education*. Boston: Houghton-Mifflin, 1978.

Abstract: This text is designed primarily for teachers in special and remedial education, but also for the support system of special education students, including counselors, educational administrators, school psychologists, and social workers. No prior knowledge of measurement and statistical concepts is assumed.

Parts 1 and 2 provide a general overview of an orientation to assessment. Part 3 provides detailed discussions and assessment of achievement, intelligence, adaptive behavior, and readiness. Part 4 is integrative and deals with the application of assessment practices in special and remedial education. Part 4 is felt to be particularly applicable to this module.

Thurlow, M. L., & Ysseldyke, J. E. Current assessment and decision-making practices in model LD programs. Learning Disability Quarterly, 1979, 2(4), 15-24.

Abstract: Data from questionnaires completed by forty-four Child Service Demonstration Centers were analyzed. Information was provided on the number of children served, the LD definition used, the kinds of assessment data collected, and the purpose for which they were used; the specific assessment devices used to collect data, and the purpose for which they were used; the typical composition of the placement team, and the major sequential steps in the assessment/decision-making process. Results suggested that assessment and decision-making in the field of learning disabilities are characterized by variability and inconsistency. The implications of the findings for current assessment practices are summarized.

Ysseldyke, J. E., & Algozzine, B. Perspectives on assessment of learning disabled students. Learning Disability Quarterly, 1979, 2(4), 3-13.

Abstract: Current critical issues in assessment of learning disabled students are described with special emphasis on logical fallacies in the assessment process. New directions in assessment are specified and discussed.

APPENDIX D

Articles

BETTY A. HARE, Department Editor

Classroom Techniques

PAUL WEHMAN
STUART SCHLEIEN

Assessment and Selection of Leisure Skills for Severely Handicapped Individuals

Leisure skills programming for severely handicapped individuals is receiving increased attention. For example, availability of leisure skill materials (Quilitch & Delongchamp, 1974; Reid, Willis, Jarman, & Brown, 1978), type of materials (Quilitch & Risley, 1973), and proximity of objects (Berkson & Davenport, 1962; Wehman, 1978) are effective techniques which require minimal trainer intervention. On the other hand, more intrusive types of teacher assistance have been demonstrated, such as reinforcement of appropriate object manipulation (Favell, 1973), modeling (Gabel, Hendrickson, & Strain, 1978; Morris & Dolker, 1974; Wehman & Marchant, 1978), and task analysis (Marchant & Wehman, 1979; Nietupski & Williams, 1974; Wehman & Marchant, 1977; Wehman, Renzaglia, Berry, Schutz, & Karan, 1978).

These efforts represent a significant step toward documenting the leisure skill capabilities of severely and profoundly handicapped individuals. Regrettably however, little information is available relevant to assessing age-appropriate skills. Furthermore, once a commitment to leisure education has been made, teachers and other practitioners are faced with the following question: *Which leisure skill(s) should be selected for instruction?* With the large number of leisure skills (e.g., games, hobbies) available, and the recreation skill deficits characteristic of most severely handicapped individuals, assessment and skill selection are processes critical to efficient instruction.

The purpose of this paper is twofold. First, leisure skill variables will be identified for assessment. These factors provide different types of data about an individual's leisure activity. Second, certain criteria will be proposed which may be used as guidelines for skill selection. The present report is intended for practitioners who work with individuals of all ages classified as moderately, severely, and profoundly mentally retarded (IQ = 0–51).

Leisure Skill Variables for Initial Assessment

Assessment of an individual's leisure skill competencies is necessary before designing an appropriate program. Initial assessment will help determine which skills the participant can perform independently and which skills require verbal, gestural, or physical assistance. Unfortunately, we are not aware of any published or unpublished leisure skill inventories or criterion-referenced curriculum guides which are sensitive to the unique needs and problems of severely and profoundly handicapped persons. Although work is underway (e.g., Wehman, 1979), at this point it is necessary to use leisure skill inventories designed for higher-functioning individuals or recreation activity guides with activities or skills that

BEST COPY AVAILABLE

have not been task analyzed. In the absence of a comprehensive curriculum guide, the variables listed below highlight important leisure skill areas for functional assessment.

Proficiency of Leisure Skill: Task Analytic Assessment

Although there are a number of areas which can be assessed in a recreation environment, an initial consideration must be *whether the individual knows how to interact with the materials.* Stated another way, when given leisure skill materials, can the participant use them appropriately? If not, then systematic instruction is required.

What is required for evaluating leisure skill proficiency is task analytic assessment (Knapczyk, 1975). An instructional objective must be written for a given material. The objective should reflect the specific skill which the teacher wants the child to learn. An example of a task analytic assessment for playing with a spinning top is provided in Table 1. This table contains an instructional objective, a task analysis for playing with a top, and the verbal cue provided during the assessment. The recording form indicates that for the first five days of assessment (baseline) the child performed a total of three, three, two, four, and four steps independently. Therefore, instruction should begin at step three in the task analysis.

There are multiple advantages to this type of observational assessment. First, the information collected about the child on a particular play skill helps the teacher to pinpoint the exact point where instruction should begin. Therefore, the child does not receive instruction on skills in which he is already proficient. Another advantage is that step-by-step individualized instruction is facilitated. Evaluation of the child's proficiency with different toys over an extended period of time will also be more objective and precise, and will be less subject to teacher bias.

Duration of Activity

If an individual has some degree of proficiency with leisure materials as Reid et al. (1978) found, then the instructional variable of interest may be the duration or length of time the participant engages in activity. Duration may be assessed by recording the amount of time spent in different activities.

TABLE 1

Task Analytic Assessment for Playing with a Top

Instructional Objective: Given a top, the child will spin the top independently a minimum of five times for three consecutive days.

Verbal Cue: "Sally, play with the top."

	M	T	W	Th	F
1. S approaches top	+	+	+	+	+
2. S places hands on top	+	+	+	+	+
3. S finds handle of top	+	+	–	+	+
4. S pushes handle down on top once	–	–	–	+	+
5. S brings handle up	–	–	–	–	–
6. S brings handle down on top twice	–	–	–	–	–
7. S brings handle up each time	–	–	–	–	–
8. S brings handle down on top three times	–	–	–	–	–
9. S brings handle up each time	–	–	–	–	–
10. S brings handle down on top four times	–	–	–	–	–
11. S brings handle up each time	–	–	–	–	–
12. S brings handle down on top five times	–	–	–	–	–
13. S brings handle up each time	–	–	–	–	–
14. S stops top from spinning	–	–	–	–	–
15. S puts top away	–	–	–	–	–

BEST COPY AVAILABLE

TABLE 2

Initial Object Assessment

Leisure Skill Object	*Minutes/Seconds Engaged with Object*	*Type of Action*
1. Pinball Machine		
2. Record player		
3. Viewfinder		
4. Lincoln Logs		
5. Ball		

The length of independent leisure activity is particularly important to assess because of its relevance to most home situations where parents cannot constantly occupy time with their handicapped child. A frequently heard request from many parents is to teach the child to play independently, thereby relieving the family of continual supervision. A careful assessment of the child's duration of leisure activity before instruction will help the teacher and parents set realistic independent leisure goals for the child. Table 2 presents a sample data collection sheet.

Discriminating Between Appropriate Versus Inappropriate Object Manipulation

Another assessment issue faced by teachers and researchers is differentiating among appropriate actions with objects versus actions which would not be considered appropriate. Several play studies have failed to address this issue (Burney, Russell, & Shores, 1977; Favell & Cannon, 1977; Wehman & Marchant, 1978). Inappropriate play actions have typically been considered those behaviors which are harmful or destructive to the child, peers, or materials. However, profoundly retarded and autistic children often exhibit high rates of repetitive self-stimulatory behavior with toys, i.e., banging, pounding, slamming, which are not necessarily harmful or destructive yet still inappropriate. Furthermore, the problem is compounded since with certain objects, banging or slamming actions may be appropriate. Many children will do unusual things with toys which *might* be considered appropriate by other observers (Goetz & Baer, 1973).

Teachers are thus faced with how to assess the qualitative nature of object manipulation. There are several ways of coping with this difficulty. The first one involves using two to three observers periodically and having these observers rate the appropriateness of the action. Objective judging provides a checks and balances system for the teacher.

A second method of assessing apropriateness of object manipulation is to identify the principal actions which a nonhandicapped child of comparable mental age might do with each object (Fredericks, Baldwin, Grove, Moore, Riggs, & Lyons, 1978). These actions may serve as guidelines.

Identifying a number of fine motor categories for object manipulation is another means of coding the qualitative nature of responses. For example, Tilton and Ottinger (1964) presented nine categories, which are self-explanatory, and which were identified after extensive observational analysis of normal, trainable retarded, and autistic children. These are repetitive manual manipulation, oral contacts, pounding, throwing, pushing or pulling, personalized toy use, manipulation of movable parts, separation of parts of toys, and combinational uses of toys.

Leisure Preference Evaluation

Assessing favorite leisure activities is an important step in initiating a recreation program. The goal is to identify which, if any, activities are preferred by the participant. By employing duration assessment, the amount of time spent with each leisure material is observed and recorded.

Latency is another measure of assessing leisure preference through presenting a small number of different materials and determining the amount of time before the participant responds. McCall (1974) has used latency as a measure of the length of time which elapsed before nonretarded infants acted on a variety of objects which were presented. Each of the objects possessed different stimulus attributes such as configural complexity of sound potential. Through measuring passage of time until a response, teachers may be able to evaluate the relative attractiveness of and preferences for certain materials with severely handicapped individuals.

BEST COPY AVAILABLE

Preference	**What skills does the client already demonstrate?**
Functioning Level	**What are the client's capabilities and educational needs?**
Physical Characteristics	**What physical characteristics does the client have or lack which may interfere with leisure skill development?**
Age-Appropriateness	**Are the skills which have been selected for instruction appropriate for nonhandicapped peers?**
Access to Materials	**Does the client have access to materials (financial resources, transportation, etc.)?**
Support of Home Environment	**What persons are available in the home or neighborhood environment to reinforce leisure skill development?**

FIGURE 1.

General Factors to Assess in Selection of Leisure Skills.

Frequency of Interactions

For many severely handicapped children, an important instructional goal is to initiate and sustain interactions with peers more frequently. A relatively common occurrence may be the presence of several handicapped children playing in isolation of each other during free play (Fredericks et al., 1978). During these play situations, the potential benefits of social interaction are not accrued.

One way of assessing social interaction is a simple count of the number of times a child initiates an interaction, receives an interaction, sustains an interaction, or terminates an interaction. Duration assessment may be used to measure the length of the interaction between peers and also between the child and adults in the room. A second means of gathering more information on social interactions is the coding of specific types of interactions. Carney, and her associates Clobuciar, Corley, Wilcox, Bigler, Fleisler, Pany, and Turner (1977) have suggested 20 social interaction skills under the four interaction categories. In addition to providing sequence, these skills may be task analyzed and the child's proficiency on selected behaviors assessed. These four categories of interaction can be employed to code the qualitative nature of the interaction (Hamre-Nietupski & Williams, 1977).

Direction of Interaction

Analyzing the direction of interactions can be helpful in assessing which individuals in the environment are reinforcing to the child. As Beveridge, Spencer, and Miller (1978) observed, child-teacher interactions occur more frequently than child-child interactions, especially with severely handicapped children. Structured intervention by an adult is usually required to increase child-child interactions (Shores, Hester, & Strain, 1976).

The direction of interactions should be assessed during home visits while observing the child playing with siblings or with neighborhood children. Data on interactions with handicapped children and with nonhandicapped peers should be recorded. This type of behavioral analysis can be revealed since most nonhandicapped children do not include handicapped children in play unless prompted and reinforced by adults (e.g., Apolloni & Cooke, 1978).

A Model for Selection of Leisure Skills

While the leisure skill variables discussed above are important in the program development process, it is equally important to review systematically criteria for skill selection. The nature of the assessment data and factors involved in Figure 1 will constitute the initial skill selection. These criteria must be carefully assessed before beginning a program. They include the participant's leisure skill preference, functioning level, physical characteristics, age-appropriateness of the skill, participant's access to materials, and the quality and support available in the home environment.

BEST COPY AVAILABLE

Figure 1 outlines a sequence of questions which the teacher should ask in reviewing skills for instruction. These guidelines will be helpful in determining the range of skills which may be selected. The proposed skill selection model is currently being field tested with severely handicapped students of all ages.

Leisure Skill Preference

The initial question to consider in determining which leisure skills to select for instruction is, "How does the individual presently spend his free time?" Stated another way, this question refers to what leisure skills are currently in the individual's repertoire.

Preference is a valuable indicator because (a) it may provide the teacher with insights as to the type or category of leisure activity that the participant enjoys, (b) the activity may be used as a reinforcer for other new leisure skills which are objectives of instruction, and (c) it allows the teacher to determine what the participant can already do. As an illustration consider the presentation of a variety of leisure materials (e.g., blocks, record player, viewfinder). The teacher could observe and record which materials were preferred, how proficient the individual was with the material, and for what period of time he engaged in its use. Similar assessments might be made with outdoor recreational equipment. The data collection sheet presented earlier in Table 2 provides a means for assessing durations.

The identification of object preferences may be helpful in determining the participant's leisure skill preference; however, it does not tell the teacher what materials and/or activities to make available. Therefore, it is critical to relate appropriate leisure activities to the individual's specific educational and/or habilitation needs.

Functioning Level and Specific Educational Needs

The individual's functioning level will affect the choice of materials and activities which should be provided for assessment. Consideration of the individual's abilities across major curriculum areas cannot be ignored. Expressive and receptive language competencies, fine and gross motor development, and social skills must be evaluated in determining which leisure skills to target for instruction.

Initial screening of the student's functioning level should reveal the behaviors in this individual's repertoire and the behaviors or component skills which make up the leisure activities that are targeted for instruction. A general parity or agreement between these two questions must be made in order for the leisure skill selection to be appropriate.

Consider the following illustration. Susan's IEP indicates that at her present performance level she is unable to attend for a period longer than three seconds, yet demonstrates competent fine motor behavior (e.g., able to grasp and pick up objects, push pull objects, squeeze, release and transfer). She is usually withdrawn from others and stays in the corner engaging in high rates of self-stimulation, i.e., twisting string or picking up scraps of paper from floor and putting them in her mouth. Susan's teacher has provided a variety of card and board games for Susan and the other students in the play area. However, Susan does not play appropriately with the games. Her approximate functioning level is not at parity with the skills required in the board games.

If the IEP committee has done a good job of initial assessment and instructional objectives have been clearly specified, then selection of leisure skills may be facilitated through interrelating IEP objectives with leisure skill goals.

Physical Characteristics

In most cases, new behavior can be developed and maintained in individuals who are functioning at low developmental levels. This is done through behavioral training techniques (e.g., Kazdin, 1975). However, the participant's physical characteristics will directly affect selection of leisure skills for instruction. Individuals with severe motor impairments, extreme spasticity, or uncontrollable seizures present additional problems in the identification of appropriate leisure skills for instruction.

Even though such physical disabilities are rarely reversible, they need not interfere with leisure skill programming (Williams, Briggs, & Williams, 1979). For example, the child with spasticity in arms and hands might enter into a game of moving a ping-pong ball back and forth through use of a head pointer. This same skill might also facilitate scanning and head pointer control on a communication board. Another illustration might involve using oversized pieces in a table game for the spastic child who is unable to use standard size materials (Marchant & Wehman, 1979). The adaptions are endless and require only a teacher's creativity and an occupational/physical therapist's knowledge of motor development (Williams et al., 1979).

Age-Appropriate Level of Skill

Another variable which must be considered in the assessment and skill selection process is how age-appropriate the skill is. The principal question to consider is, "Would a nonhandicapped individual of comparable chronological age engage in this activity during free time?" Severely handicapped adults on the floor pushing a toy truck around or playing with a dollhouse are examples of inappropriate skill selection.

There is, unfortunately, limited empirical research on leisure skill programming with severely handicapped adolescents and adults. It would appear, however, that the answer to the age-appropriate problems can be found in a detailed breakdown of the skill into very small behaviors. For example, the prospect of teaching plant care to a severely or profoundly handicapped individual may appear remote. However, if plant care is divided into several skills (e.g., putting dirt in pot, putting flower in dirt, etc.) and each skill is task analyzed, then learning problems will be reduced.

Access to Materials and Events

The most capable individual will have difficulty engaging in a variety of leisure activities without access to materials or events. At the least, transportation and some degree of financial resources are involved. The following question must be considered, "Can the participant get to community events, and if unemployed, does he or she have the money to make necessary purchases?" Although leisure activities can be engaged in without money (i.e., building a snowman), usually some funds are necessary for the new materials and replenishing old materials.

There are other factors to consider as well. For example, initiating a social encounter may be difficult without knowledge of how to use a telephone (Nietupski & Williams, 1974) or if no phone is available. Interaction with toys or other play objects at home will be difficult if they are stored away in a closet or have been destroyed with no funds to purchase new ones. Similarly, many residents in state facilities have difficulty operating the television or stereo which is placed eight to 10 feet above the floor.

In short, a careful assessment of what to teach must include a look at the amount and type of materials available, the proximity and physical design of local recreational facilities, the ease of transportation, and the availability of skilled recreational personnel to provide training. An analysis of these variables will, at a minimum, facilitate a decision concerning how broad a program to establish. It will also help identify what areas need more adaption and planning.

Home Environment

Perhaps the most critical factor in leisure skill selection is evaluation of the home and neighborhood environment. The age of the individual's parents, the presence of siblings or other relatives in the home, the type of home, and the attitude of other home members will greatly influence the variety and independence of leisure activities.

Location of the home will also affect the selection process. Urban living presents different problems than sparsely populated rural areas. Sensitivity of local communities and neighborhood members to handicapped persons will also be reflected in the amount of funds which are appropriated for therapeutic recreation programming. In Table 3 is a checklist of factors to consider in evaluating the home environment for leisure skill selection.

The willingness of parents and other family members to follow through on school training programs is important as well. Marchant and Wehman (1979) found that demonstration and behavior rehearsal with a foster mother of a severely retarded child was instrumental

TABLE 3

Client Home Environment

The participant lives with ______________________________

Relationships and ages of other persons in home ______________________________

____

____

How does the participant utilize his free time at home?

Does the participant utilize his leisure time independently or with assistance in the home?

What leisure activities and interests are enjoyed by others in home?

How many hours per day do housemates/siblings spend with participant?

Do housemates have any natural (e.g. athletics, musical ability) or learned (e.g. cooking, tools) talents?

How is the participant perceived by himself and by family members?

What are the present attitudes toward recreation and leisure held by the participant's family members?

Does the participant reside in an urban or rural area?

What is the attitude of the neighborhood toward integration of handicapped persons into community programs?

Does the participant interact with others in the neighborhood?

in generalizing table game skills from the classroom to the home. Parent/professional partnership is vital to maintenance of leisure activity repertoire in severely handicapped individuals.

Summary

The purpose of this paper has been to describe several types of leisure skill competency areas which can be assessed in severely handicapped individuals. These included the *proficiency* with which objects or materials were engaged, the *length* of self-initiated action, materials *preference* by clients, and frequency and direction of *social interactions*.

In the second half of the article, guidelines for selecting leisure skills were presented. A variety of areas were identified as critical to the skill selection process. Client preference for different materials, functioning level, age-appropriateness of activity, and support of the home environment were among the principal criteria clusters.

When these assessment and skill selection guidelines are provided in conjunction with logically sequenced recreation curriculum and instructional technology, the application of the systematic instruction process (Wehman & McLaughlin, in press) to leisure skill development is complete. What remains is the continued development, field-testing, and validation of leisure skill curricula. The items sequenced in these curricula will then serve as appropriate criterion-referenced skills for assessment.

References

Apolloni, T., & Cooke, T. Integrated programming at the infant, toddler, and preschool levels. In M. Guralnick (Ed.), *Early intervention and the integration of handicapped and nonhandicapped children.* Baltimore: University Park Press, 1978.

Berkson, G., & Davenport, R. K. Stereotyped movement of mental defectives. I. Initial survey. *American Journal of Mental Deficiency*, 1962, *66*, 849–852.

Beveridge, M., Spencer, J., & Miller, P. Language and social behavior in severely educationally subnormal children. *British Journal of Social and Clinical Psychology*, 1978, *17(1)*, 75–83.

Burney, J., Russell, B., & Shores, R. Developing social responses in two profoundly retarded children. *AAESPH Review*, 1977, *2*, 53–63.

Carney, I., Clobuciar, A., Corley, E., Wilcox, B., Bigler, J., Fleisler, L., Pany, D., & Turner, P. Social interaction in severely handicapped students. In *The Severely and Profoundly Handicapped Child.* Springfield, Illinois: State Department of Education, 1977.

BEST COPY AVAILABLE

Favell, J. Reduction of stereotypes by reinforcement of toy play. *Mental Retardation*, 1973, *11(4)*, 24-27.

Favell, J., & Cannon, P. R. Evaluation of entertainment materials for severely retarded persons. *American Journal of Mental Deficiency*, 1977, *8*, 357-361.

Fredericks, H. D., Baldwin, V., Grove, D., Moore, W., Riggs, C., & Lyons, B. Integrating the moderately and severely handicapped preschool child into a normal day care setting. In M. Guralnick (Ed.), *Early intervention and the integration of handicapped and nonhandicapped children*. Baltimore: University Park Press, 1978.

Gable, R., Hendrickson, J., & Strain, P. S. Assessment, modification, and generalization of social interaction among severely retarded multihandicapped children. *Education and Training of the Mentally Retarded*, 1978, *13*, 279-286.

Goetz, E., & Baer, D. Social control of form diversity and the emergence of new forms in children's blockbuilding. *Journal of Applied Behavior Analysis*, 1973, *6*, 209-217.

Hamre-Nietupski, S., & Williams, W. W. Implementing of selected sex education and social skills programs with severely handicapped students. *Education and Training of the Mentally Retarded*, 1977, *12*, 364-372.

Kazdin, A. E. *Behavior modification in applied settings*. Homewood, Ill.: Dorsey Press, 1975.

Knapczyk, D. Task analytic assessment of severe learning problems. *Education and Training of the Mentally Retarded*, 1975, *10*, 74-77.

Marchant, J., & Wehman, P. Teaching table games to severely retarded children. *Mental Retardation*, 1979, *17*, 150-152.

McCall, R. Exploratory manipulation and play in the human infant. Monographs of the Society for Research in Child Development, 1974, 39(2).

Morris, R., & Dolker, M. Developing cooperative play in socially withdrawn retarded children. *Mental Retardation*, 1974, *12(6)*, 24-27.

Nietupski, J., & Williams, W. W. Teaching severely handicapped students to use the telephone to initiate selected recreational activities and to respond appropriately to telephone requests to engage in selected recreational activities. In L. Brown, W. Williams, & T. Crowner (Eds.), *A collection of papers and programs related to public school services for severely handicapped students*. Madison, WI: Madison Public Schools, 1974.

Quilitch, H. R., & Delongchamp, G. D. Increasing recreational participation of institutional neuropsychiatric residents. *Therapeutic Recreation*, 1974, *8*, 56-57.

Quilitch, H. R., & Risley, T. The effects of play materials on social play. *Journal of Applied Behavior Analysis*, 1973, *6*, 573-578.

Reid, D., Willis, B., Jarman, P., & Brown, K. Increasing leisure activity of physically disabled retarded persons through modifying resource availability. *AAESPH Review*, 1978, *3*, 78-93.

Shores, R., Hester, P., & Strain, P. S. The effects of amount and type of teacher-child interaction during free play. *Psychology in the Schools*, 1976, *13*, 171-175.

Tilton, J., & Ottinger, D. Comparison of toy play behavior of autistic, retarded, and normal children. *Psychological Reports*, 1964, *15*, 967-975.

Wehman, P. Effects of different environmental conditions on leisure time activity of the severely and profoundly handicapped. *Journal of Special Education*, 1978, *12*, 183-193.

Wehman, P. Toward a recreation curriculum for developmentally disabled persons. In P. Wehman (Ed.), *Recreation programming for developmentally disabled persons*. Baltimore: University Park Press, 1979.

Wehman, P., & Marchant, J. Developing gross motor recreation skills in children with severe behavioral handicaps. *Therapeutic Recreation Journal*, 1977, *11(2)*, 48-54.

Wehman, P., & Marchant, J. Improving free play skills of severely retarded children. *American Journal of Occupational Therapy*, 1978, *32(2)*, 100-104.

Wehman, P., & McLaughlin, P. J. *Program Development in Special Education: Designing Individualized Education Programs*. New York: McGraw-Hill, in press.

Wehman, P., Renzaglia, A., Berry, G., Schutz, R., & Karan, O. C. Developing a leisure skill repertoire in severely and profoundly handicapped adolescents and adults. *AAESPH Review*, 1978, *3*, 162-172.

Williams, B., Briggs, N., & Williams, R. Selecting, adapting and understanding toys and recreation materials. In P. Wehman (Ed.), *Recreation Programming for Developmentally Disabled Persons*. Baltimore: University Park Press, 1979, P. 15-36.

PAUL WEHMAN *is Assistant Professor, Division of Educational Services, and* **STUART SCHLEIEN** *is Coordinator of Leisure Skills Curriculum Project, both of Virginia Commonwealth University, Richmond, Va.*

The development and dissemination of this paper was partially supported by a State Developmental Disabilities Grant, Department of Mental Health and Mental Retardation, Richmond, Virginia.

The skill selection criteria guidelines were stimulated through thoughtful interactions with Wayne Fox.

BEST COPY AVAILABLE

INFORMAL ASSESSMENT OF WRITTEN EXPRESSION

James A. Poteet

***Abstract*. This article presents an approach to informal assessment of written expression which allows the teacher to determine (a) the skills acquired and correctly used by the student, (b) the skills needing remediation, and (c) the skills to be introduced. A *Checklist of Written Expression Skills* is provided as an aid in this informal assessment procedure. Results of the assessment are directly related to the instructional objectives and annual goals in the student's IEP. Written expression and informal assessment are defined and development of the language system is discussed. Data required by the teacher prior to informal assessment are described as (a) the broad spectrum of written expression skills, (b) requirements of the specific writing task being assessed, and (c) student characteristics related to the student's language system.**

Informal assessment is an invaluable aid for a learning disabilities teacher both in daily teaching and in providing diagnostic information as a member of the evaluation team. Results of informal assessment assist the teacher in deciding what specific skills to teach. This article will focus on the informal assessment of written expression.

Written expression can be defined as a visible representation of thoughts, feelings, and ideas using symbols of the writer's language system for the purpose of communication or recording.

A person's language system develops hierarchically beginning with inner language acquired by the association of a variety of experiences with words (Myklebust, 1965). Maturity in language development is achieved by the sequential development of (a) ideas through experiences and thought, (b) comprehension of the spoken language of others through listening, (c) oral expression of ideas to others by use of spoken language, (d) reading of the written language, and finally (e) written expression. In other words, expression is the last stage of the language system to develop; it is highly complex and requires the integration of all previous stages (Myklebust, 1968).

Prior to informal assessment of written expression, the teacher must possess the following information:

1. broad spectrum of written expression skills,
2. requirements of the specific writing task, and
3. characteristics of the student being assessed as they relate to the developmental nature of the student's language system.

JAMES A. POTEET, Ph.D., is Associate Professor of Special Education, Ball State University, Muncie, IN.

The broad spectrum of written expression skills. The teacher must know the scope of necessary skills which comprise written expression (see Figure 1) as well as the sequence in which they are taught. By comparing samples of the student's writing with a scope-and-sequence chart, the teacher can determine which skills the student has and has not yet acquired. Knowing the sequence in which writing skills should be taught, the teacher can determine which unknown skills should be introduced first and how to prioritize the skills that need remediation.

The language arts textbook in use by the local school district is usually the best source for such a scope and sequence of skills. Possibly, a local curriculum guide has been developed which gives directions for skills development by grade level. If such materials are not available, an excellent scope-and-sequence chart for grades one through eight is provided in Hammill and Bartel (1978, pp. 186-191), and for grades one through twelve in Poteet (1980).

By viewing the spectrum of written expression as being comprised of components, each with unique elements, the teacher can more easily structure both assessment and instructional practices. While a variety of components have been conceptualized (McCrimmon, 1973; Hodges & Whitten, 1977), the following components and their elements allow the teacher to approach informal assessment in a systematic way:

I. Penmanship
II. Spelling
III. Grammar
IV. Ideation

I. *Penmanship* concerns the legibility of writing. Handwriting quality is the focus in this component. The spacing of the writing on the total page, the spacing of sentences, words, and letters, slant, letter formations, pressure on the paper, and pencil grip are important elements of penmanship quality to be considered.

II. *Spelling*. Words incorrectly spelled and nonrecognizable "words" are considered under this component. Error patterns might be characteristic of the misspellings and will be discussed later in the informal assessment discussion. The percent of incorrectly spelled words can guide the teacher in determining if a remediation program is necessary for the student being assessed. Elements to consider under this component are miscalled rules, letter insertions, omissions and substitutions, phonetic spelling, directional confusion, schwa and r-controlled vowels, letter orientation, and sequence.

III. *Grammar* deals with the accepted rules and conventions of written expression as seen in contemporary American-English. It covers the mechanical, semantic, and syntactical aspects of our written language and is divided into: a. Capitalization, b. Punctuation, and c. Syntax, each with its own subelements.

Subelements under Capitalization specify proper usage of capital letters as proper nouns, proper adjectives, first word in a sentence, first word in a line of verse, first word in a quotation, principal words in a literary title, personal titles, the use of "I" and "O", personification, letter salutation, complimentary ending in a letter, and possible other usages. Subelements under punctuation include uses of the period, comma, apostrophe, quotation marks, question mark, semicolon, exclamation mark, colon, the dash, parentheses, brackets, and the slash. Syntax, which concerns the relationships among words, has the following subelements: parts of speech (verbs: transitive, intransitive, active, passive, past, present, future, infinitives, gerunds as nouns, and participles as adjectives; pronouns; adjectives: articles, modifiers, predicate adjectives; adverbs; prepositions; conjunctions: coordinating, subordinating; interjections), agreement of subject and predicate, correct use of nominative, possessive, and objective case, pronoun reference, order and position of words and phrases (dangling modifiers), parallelism, the use of abbreviations and/or numbers, and the paragraph (cohesiveness, topic sentence, transition sentence).

IV. *Ideation* concerns the thoughts and ideas the writer has at command while writing. The writer must have some idea to express or to convey. The content can reflect varying levels of abstract thinking, creativity, and emotional quality. Word choice can render the writing from trite to spellbinding. Effectiveness often hinges on the productivity

Figure 1

CHECKLIST OF WRITTEN EXPRESSION*

Student's Name______________Grade Placement_____Teacher___________Grade_____

	TA	A	I	R	Notes
I. PENMANSHIP					
Rating: 1 2 3 4 5					
A. Spacing on the page					
B. Spacing of the sentences					
C. Spacing of the words					
D. Spacing of letters					
E . Slant					
F. Letter formations					
G. Pressure on the paper					
H. Pencil grip					
II. SPELLING					
_____% misspelled					
A. Miscalled rule					
B. Letter insertion					
C. Letter omission					
D. Letter substitution					
E. Phonetic spelling					
F. Directional confusion					
G. Schwa or r-controlled vowels					
H. Letter orientation					
I. Sequence					
J. Other					

Figure 1

CHECKLIST OF WRITTEN EXPRESSION (Continued)

III. GRAMMAR	TA	A	I	R	Notes
A. *Capitalization*					
1. proper noun					
2. proper adjective					
3. first word in a sentence					
4. first word in a line of verse					
5. first word in a quotation					
6. principal words in a title					
7. personal title					
8. use of "I" or "O"					
9. personification					
10. salutation in a letter					
11. complimentary close in a letter					
12. other					
B. *Punctuation*					
1. period					
2. comma					
3. apostrophe					
4. quotation marks					
5. question mark					
6. semicolon					
7. exclamation mark					
8. colon					
9. the dash					
10. parentheses					
11. brackets					
12. the slash					
C. *Syntax*					
1. parts of speech					
a. verbs					
b. nouns					
c. pronouns					
d. adjectives					
e. adverbs					
f. prepositions					
g. conjunctions					
h. interjections					
2. agreement					
3. case					
4. pronoun reference					
5. order/position of words					
6. parallelism					
7. abbreviations/numbers					
8. the paragraph					

Figure 1
CHECKLIST OF WRITTEN EXPRESSION (Continued)

IV. IDEATION

A. *Type of writing*

1. story_____ 2. poem_____ 3. letter_____ 4. report_____ 5. review_____

B. *Substance*

1. Naming_____ 2. Description_____ 3. Plot_____ 4. Issue_____

C. *Productivity*

1. Number of words written_____ 2. Acceptable number_____ 3. Too few_____

D. *Comprehensibility*

Easy to understand_____ Difficult to understand_____ Cannot understand____

____perseveration of words ____illogical

____perseveration of ideas ____disorganized

E. *Reality*

_____Accurate perception of stimulus or task

_____Inaccurate perception of stimulus or task

F. *Style*

1. Sentence Sense
   a. Completeness — Tallies:
      (1). complete sentences ________________
      (2). run-on sentences ________________
      (3). sentence fragments________________
   b. Structure
      (1). simple ________________
      (2). compound ________________
      (3). complex ________________
      (4). compound/complex________________
   c. Types
      (1). declarative ________________
      (2). interrogative ________________
      (3). imperative ________________
      (4). exclamatory ________________
2. Tone
   a. intimate_____ b. friendly_____ c. impersonal_____
3. Word Choice (N = none, F = few, S = some, M = many)
   a. formality
      formal_____ informal_____ colloquial_____
   b. complexity
      simple_____ multisyllable_____ contractions_____
   c. descriptiveness
      vague_____ vivid_____ figures of speech_____
   d. appropriateness
      inexact words_____ superfluous/repetitions_____ omissions_____

---

*Adapted from *The Inventory of Written Expression and Spelling* (Poteet, 1980).

level of the writing; are enough words used to adequately represent the ideation underlying the content? Elements within this component to consider are: a. Type of Writing, b. Substance, c. Productivity, d. Comprehensibility, e. Reality, and f. Style.

Five major types of writing are taught in the schools each with its own requirements for effectiveness. They are stories, poems, letters, reports, and reviews. Stories are the most familiar to students and are probably the easiest of all types of writing to produce. Often story starters such as pictures or a topic sentence are used as motivators. Poems as well as stories require creative skills, and they are often structured according to topics or style. Writing of business and friendly letters is taught in the elementary grades. Reports which make use of factual information can range from a simple telling about a vacation trip to a detailed scientific or literary report often assigned in the upper grades. Reviews consist of personal opinions about books, movies, TV shows, plays, or musical performances.

Substance concerns the abstract quality of the writing which can easily be categorized into the following four abstraction levels: (1) Naming, (2) Description, (3) Plot, or (4) Issue (Poteet, 1980). If the writing task is to write a story about a picture, then in level one, the writer essentially lists objects or people in the picture ("I see a fire engine, a ball, a house, a man, a tree, a car"). Such writing is analogous to oral expression of labeling, a rather immature form of communication. Level two, description, is illustrated by an objective reporting of the content of a picture ("I see a man. The grass is green. The bus is blue. The girl is running. It is raining."). Level three which represents a plot is a completed story about the picture however short or poorly expressed. Level four goes beyond a simple plot to deal with some moral theme or issue.

Productivity, the number of words written, is an indicator of maturity in written expression with more words written as the student grows older. Hermreck (1979) found that non-handicapped students writing a story about a picture selected by the teacher, wrote an average of 83 words in grade three, 94 words in grade four, 141 words in grade five, and 200 words in grade six. She also found that learning disabled students wrote an average of 40 words in grade three, 54 words in grade four, 70 words in grade five, and 170 words in grade six. Consequently, it seems that learning disabled students write fewer words than their non-handicapped peers (Hermreck, 1979; Myklebust, 1973; Poteet, 1979). Some writers simply do not use enough words to give the reader a feeling of completeness about what is read. In some instances, words are obviously omitted and in other instances superfluous words distract from the continuity of the writing.

Comprehensibility is concerned with the ease of understanding what the student wrote. *Reality* assesses the accuracy of the writing which must reflect a realistic perception of the stimulus or of the writing task. *Style* is the unique way a piece of writing is written. It is highly related to the purpose of the writing (McCrimmon, 1973). For writing to be effective, correct sentence sense, appropriate tone, and appropriate choice of words are necessary; these are the elements of style.

Sentence sense subelements to consider are: a. completeness as observed in complete sentences, run-on sentences, and sentence fragments; b. structure as observed in simple, compound, complex, and compound/complex sentences; and c. types of sentences as observed in declarative, interrogative, imperative, and exclamatory sentences. The *tone* of the writing can be described in relationship to the distance the writer establishes between himself/herself and the reader. The writing can be described as intimate, friendly, or impersonal.

The emotional quality of writing is best assessed by inspecting *word choice*, sometimes called "diction". Words can be analyzed in terms of formality, complexity, descriptiveness, and appropriateness. Words have been described as formal ("to depart"), informal ("to leave"), and colloquial ("to split"). Some words can be quite complex consisting of several syllables while others are simple, short, or contractions. Descriptiveness of the words can range from concrete (Benny Goodman) to abstract (musician).

Descriptive words can relate to the senses of touch (sandy), taste (salty), smell (pungent), sound (hiss), or sight (flicker) (McCrimmon, 1973). Some writers use such vague words and trite expressions that the reader is left with uncertainty about the writer's intent. These writers use words such as "all that stuff", "and all", and "everything like that" Words may also be vivid and exciting; they also may illustrate figures of speech. Appropriateness of word choice demands that the exact words or idioms be used (not "I *except* your invitation"), that superfluous words and useless repetitions not be used, and that necessary words not be omitted.

When all of the above four major components and their elements have been analyzed in a written expression sample, a *comprehensive* assessment of the broad spectrum of writing skills will have been conducted. Comprehensive informal assessment must be viewed both as (a) holistic, all components are considered as integrated into one complete whole, and as (b) atomistic, specific skills or elements of the components are considered as the bits and pieces which are combined to construct the final product.

It is not necessary to employ *all* elements of the assessment for every student being assessed. The more advanced skills understandably have not been acquired by the younger students and should not be considered during assessment. The severity of the student's learning disability, the skills which have been taught, and the curriculum appropriate to the student's grade placement must be considered when assessment is conducted. The earlier skills are of primary interest for the younger or the more disabled student while the more advanced skills would be appropriate for the adolescent and young adult or the less disabled student. For the teachers who might need to review the complexities of written expression, the *Harbrace College Handbook* (Hodges & Whitten, 1977) offers an excellent overview.

The requirements of the specific writing task. The teacher must consider the skills necessary and appropriate to the specific writing task that is being assessed. While capitalization and punctuation skills are necessary across all *types* of writing, word choice and sentence structure would logically be different in a friendly letter than in a formal report or would even differ between types of stories. By considering the task requirements before assessment, the teacher can establish appropriate and fair expectations of skill use. Conversely, skills that are not appropriate can be identified and not given consideration during assessment. Familiarity with the broad spectrum of skills discussed above facilitates this process.

The characteristics of the student. In light of the development of the language system discussed earlier, the teacher must be familiar with the student's characteristics as they relate to this development. Knowledge of the student's personal and educational experiences and interests can be used to build reasons for writing. Hearing acuity and its related listening comprehension skills must be known as well as the student's speech patterns and use of oral expression. Dialects and cultural influences affect the way students talk and often how they write.

The teacher must be familiar with the student's oral expression. Many times, knowing how a student talks helps determine if the error in written expression should receive high or low priority for remediation. For instance, if the student wrote "Are hole famly was washing TV one night", one would seriously doubt the literal interpretation of the sentence. Probably, the student actually says "watching". If so, the error suggests a recall error of graphically representing the /tch/ sound. Asking the student to read aloud what was written can also provide clues to the reasons for the errors.

Since one cannot write what one cannot read, a knowledge of reading skills is important. Knowledge of fine-motor skills can explain the observed quality of the student's penmanship. Knowledge of the student's cognitive development can explain the level of ideation content in the student's writing.

An Approach to Informal Assessment in Written Expression

Informal assessment is that set of procedures designed to provide the teacher or diagnostician with precise answers to specific questions not available from standardized, norm-referenced tests. The purpose of such assessment is to determine appropriate instructional strategies for the student being assessed. Informal procedures are not standardized or norm referenced, usually do not provide a score, may be published or teacher made, and usually describe the behavior rather than quantifying it. These procedures also may provide broad general information as obtained from an informal reading inventory or may pinpoint precise skills and error patterns suggested for remediation.

While Hammill and Bartel (1978) view informal assessment as "by far the most profitable assessment procedure", they caution that the method is only as good as the teacher's competency, that the reliability of the teacher as an observer is always unknown, and that the teacher's experience may not substitute for normative data.

The use of informal assessment, however, does allow the teacher to determine what to teach and how to teach it. When using informal assessment the teacher must always consider the student's educational history as well as the student's past and present curriculum, and the student's current grade placement. To decide *what* to teach the teacher must first determine:

(a) the skills acquired and correctly used by the student
(b) the skills needing remediation
(c) the skills to be introduced

Skills acquired and correctly used can be determined by analyzing several samples of the student's writing over a period of a few weeks to insure that the student reliably displays the observed skills. A variety of types of writing, such as letters, reports, etc., as appropriate for the student's grade placement, should be inspected since the use of skills varies with the type of writing produced. The requirements of the writing task, then, will change the focus of the diagnostic inquiry.

Skills needing remediation and skills to be introduced are determined by noting errors and error patterns in the writing samples. Conceivably, any error noted on the student's sample could readily be designated for remediation. A story written by a bright third grade student about an intriguing mystery with much dialogue might indicate that she does not make use of quotation marks. By reference to a scope-and-sequence chart the teacher notices that quotation marks are typically taught at the fourth grade level. This skill then would not be considered *remedial* for this third grade student, but with an individualized language arts program, the teacher could *introduce* this skill especially since the student uses dialogue in her writing.

The skill to introduce might be a more advanced skill in the sequence of skills currently being used by the student; or it might be a "lateral expansion" of some current skill being correctly used. For instance, the use of the exclamation mark might be introduced to give variety to the type of sentences written and to generate excitement in the thematic material ("Wow! Would you believe it!?!"). A lateral expansion of some current skill might deal with word choice, not so much to develop a new vocabulary, but to use existing word knowledge to create sensory awareness and reader involvement ("The green slime oozed between his toes as he felt the prickly weeds cutting his legs").

The decision as to which skill to remediate or which skill to introduce must rest with the teacher. To assist in this decision making, knowledge of the scope and sequence of written expression skills, the instructional program available to the student, the student's aspirations, previous skills taught in the curriculum, the requirements of the writing task, and skills appropriate for the student's current grade placement must all be considered.

For the purposes here, the following approach to informal assessment is designed for use with a story written by a student after being assigned by the teacher using some stimulus as a story starter. Notations are made directly on the written sample for use by the teacher. The page of writing probably should not be returned to the student; a large number of marks could be overwhelming.

The approach to analyzing the written sample will follow the sequence outlined in the *Checklist of Written Expression* (see Figure 1). Beside each entry in the checklist, place a check mark in one of the four columns which best describes the student's achievement in written expression in the first three components. The far left column (labeled "TA") is for those skills or characteristics which the teacher decides are *too advanced* for consideration at this time or are *not appropriate* to the writing task being assessed. The next column (labeled "A") is for skills that are *adequately* used in the sample. Skills checked in these two columns do not require the teacher's attention for instruction at this time. The next column (labeled "I") is for those skills which need to be *introduced* to improve the quality of the student's writing. Typically, they are not observed in the student's writing. The last column (labeled "R") is for those skills which need *remediation* or *review*. These skills are used either incorrectly or inconsistently. Skills checked in these last two columns require the teacher's immediate attention for instruction. To the right of the four columns is space for writing specific errors which need to be remediated or for brief notes regarding instructional planning.

Before assessing the student's written sample, the teacher might scan the checklist and determine those skills which are "too advanced for consideration at this time" or "not appropriate to the writing task" and place a check mark in the column labeled "TA" beside those skills. All remaining entries on the checklist should be inspected in the written sample.

Once the student has completed the writing task the teacher should ask the student to read aloud what was written and should note any deviance from the written sample (Zigmond, 1976). This procedure helps the teacher deciphering misspelled and unrecognizable "words". Now the teacher should begin the careful and critical analysis of the written sample, component by component, using the checklist.

I. *PENMANSHIP*. First, obtain your overall impression of the quality of handwriting. Rate it from 1 (It's a mess!) to 5 (Beautiful!) and circle the rating number found on the checklist. Then, notice the use of space on the page, spacing of sentences, spacing between words, and finally spacing between letters. Next, determine if the slant is appropriate. Then look at the letter formations, but always in light of the penmanship curriculum. Notice if the poor quality of formations suggests a combined use of upper- and lower-case letters. Then analyze the types of letters showing poor quality. For instance, inspect the short round letters that begin alike (a, c, o, etc.), the tall letters (l, k, t, etc.), letters easily reversed or rotated (p, b, d, q, m, w, n, u), letters that go below the line (g, j, p, q, y), and slanting letters (v, w). Other patterns might be equally important depending on whether manuscript or cursive is used. Write the specific letters needing remediation on the checklist in the space provided.

For those who like visible examples, the *Zaner-Blose Evaluation Scales* can provide samples of varying degrees of quality for grades one through eight which can be used for comparison, or even for scoring, purposes. However, the best comparison is with the student's *classmates*. Collect from all students a sample of their best writing of the same sentence. Divide the samples into five stacks rated "poor", "fair", "average", "very good", and "excellent". Divide each of the five stacks into five stacks similarly classified. Continue in this manner and eventually choose the best representative of the "average" classification in each of the original five categories. With these five representative samples from each classroom, an individual student's written expression sample can easily be compared and judged as "poor" to "excellent". Students whose overall quality is ranked as "poor" or "fair" on the checklist should receive a remediation program for penmanship skills.

II. *SPELLING*. Underline all misspelled words on the written sample and determine the percentage of misspelled words by dividing the number of misspelled words by the total number of words written. Decide at what percent of misspelled words an intensive spelling remediation program should be implemented. Analyze the misspelled words to determine if some sort of error *pattern* is consistent. Many spelling errors result from confused recall of spelling rules. Others are related to insertion of unnecessary letters (umberella for umbrella), omission of necessary letters (famly for family), substitution of letters (mush for must), phonetic spelling (sed for said), directional confusion (was for saw), schwa and r-controlled vowels (tabel for table; doller for dollar), orientation confusion of letters (d,b,p,q,n,u,m,w), sequence (aminals for animals). In the space provided on the checklist, tally the types of errors made in the misspelled words so that the frequency of the errors can be discerned. Then determine which error patterns need remediation and check the remediation column beside those errors. Write those words that are frequently used and misspelled as they should receive immediate attention.

III. *GRAMMAR*. A. *Capitalization*. Capitalization errors are easy to spot. Circle the errors on the written sample. Capital letters are either omitted or inserted where they should not be used. In the space on the checklist write "om" for omission or "in" for insertion beside the entries *each* time an error is observed. This information will show the teacher the types of capitalization errors made and the frequency of such errors. The appropriate column should then be checked. B. *Punctuation*. Follow the same procedure with punctuation errors as was used with capitalization errors. Determining capitalization error is relatively straightforward, but determining punctuation errors is less precise. There is a wide divergence of opinion regarding the proper place for a comma. The teacher can only inquire about the most current use that is part of the student's contemporary curriculum. C. *Syntax*. To judge the written sample in terms of syntax, inspect the use of each of the subelements listed on the checklist. Determine if the appropriate parts of speech are used and used correctly. Is there agreement between the subject and the verb in each sentence? Is the appropriate case used? Are pronoun references clear? Do dangling modifiers or unusual word order detract from the meaning of the writing? Has parallelism been effectively used? Are abbreviations and numbers inserted incorrectly within sentences? Does the paragraph contain a major idea and is it cohesive? Do the paragraphs have a topic sentence — not necessarily the first sentence? Do they have transitional sentences? Do they flow logically in the correct sequence? Does the writing have an introduction and some type of logical conclusion?

IV. *IDEATION*. A. *Type of writing*. Check the entry on the checklist that describes the type of writing being assessed. B. *Substance*. If the type of writing is a story, check the level of substance reflected in the writing, otherwise omit this section. C. *Productivity*. Simply count all the words written and record the total on the checklist. Check if the number of words written are appropriate for the student's age and grade placement. D. *Comprehensibility*. Is the writing easy or difficult to understand? If it is difficult to understand or if it cannot be understood, describe the difficulty in the space provided after checking any descriptor on the checklist that might be appropriate. E. *Reality*. Does the writing reflect an accurate perception of the stimulus or of the writing task/assignment? F. *Style*. (a) *Sentence sense*. After reading *each* sentence, place a tally mark in *each* of the three sections (a. Completeness, b. Structure, and c. Types) on the checklist under Sentence Sense to describe that sentence. For instance, one sentence may be checked as complete, simple, and declarative, while another might be checked as complete, compound, and interrogative. A tally of those descriptors for each sentence reveals a visible record of the variety and types of sentences used. With this information, the teacher can more easily decide those aspects of sentences which need to be remediated or introduced. Since the four columns are not used in the Ideation component, circle those elements which should be remediated and write "Introduce" beside those to be introduced to the student. (b). *Tone*. Determine the level of the tone used in the writing and note it with a check mark on the checklist. Most writing is described as friendly. An intimate tone makes use of frequent personal pronouns while an impersonal tone is found in formal reports. (c). *Word choice*. Notice the types of words the student used in the written sample. Word choice is strongly influenced by the type and purpose of the writing. Designate the quantity of the types of words used in the writing by marking "N" for none, "F" for few, "S" for several, and "M" for many beside the descriptors on the checklist. There are three descriptors beside each of the subelements of formality, complexity, descriptiveness, and appropriateness. Circle the descriptors which suggest the need for instructional attention.

For instance, a teacher might put an "N" beside the descriptor *vivid* in a rather dull piece of writing and circle it as a reminder to teach the use of more colorful words.

Where Do You Go from here?

By now the teacher should have a comprehensive analysis of the student's written-sample. Specific skills to review and to remediate are checked and new skills to introduce have been determined. How does the teacher prioritize the skills which have been selected for instruction? While each teacher will have his/her own preferences, one approach to determining which skill to teach first can be found by reference to a scope-and-sequence chart. Skills taught in grade one might receive priority over skills taught in grade two. For the adolescent the priority skills should relate directly to writing skills required for career opportunities and daily living such as job application, check writing, test taking, outlining, report writing, etc. (see Alley & Deshler, 1979).

Given the set of specific skills targeted for remediation and introduction, the teacher should list them by priority in sequence for instructional planning. The most important skills should be taught first, and they should be the major instructional objectives on the student's individualized education program (IEP).

Informal assessment conducted by professional educators will yield results which suggest instructional strategies that are more educationally relevant and instructionally appropriate than strategies selected by a variety of other diagnostic approaches. When you consider that the student's teacher has the primary responsibility for instruction, shouldn't the teacher also have the primary responsibility for determining *what* is to be taught?

REFERENCES

Alley, G., & Deshler, D. *Teaching the learning disabled adolescent: Strategies and methods.* Denver: Love Publishing Company, 1979.

Hammill, D.D., & Bartel, N.R. *Teaching children with learning and behavior problems* (2nd ed.). Boston: Allyn & Bacon, 1978.

Hermreck, L. *A comparison of the written language of LD and non-LD elementary children using the Inventory of Written Expression and Spelling.* Unpublished master's thesis. University of Kansas, 1979.

Hodges, J.C., & Whitten, M.E. *Harbrace college handbook* (8th ed.). New York: Harcourt, Brace & Jovanovich, 1977.

McCrimmon, J.M. *Writing with a purpose.* Boston: Houghton-Mifflin, 1973.

Myklebust, H.R. *Development and disorders of written language. Volume one. Picture Story Language Test.* New York: Grune & Stratton, 1965.

Myklebust, H.R. *Progress in learning disabilities. Volume One.* New York: Grune & Stratton, 1968.

Myklebust, H.R. *Development and disorders of written language. Volume two. Studies of normal and exceptional children.* New York: Grune & Stratton, 1973.

Poteet, J.A. Characteristics of written expression of learning disabled and non-learning disabled elementary school students. *Diagnostique,* 1979, 4(1), Winter/Spring, 60-74.

Poteet, J.A. *The Inventory of Written Expression and Spelling.* Unpublished manuscript, Ball State University, 1980.

Zigmond, N. *Teaching children with special needs.* Dubuque, IA: William C. Brown, 1976.

Requests for reprints should be addressed to James Poteet, Special Education, Ball State University, Muncie, IN 47306.

BEST COPY AVAILABLE

STEVEN R. FORNESS
ROBERT L. THORNTON
AMY A. HORTON

Assessment of Applied Academic and Social Skills

Abstract: Assessment of relevant academic and social skills of developmentally disabled adolescents or young adults would seem to be most effective if done in the context of an applied or prevocational setting. An assessment instrument was developed to measure applied reading and comprehension skills, direction following (using information from the reading passages), applied number skills, and social adaptation to a work situation. Comparison of the scores of 30 developmentally disabled adolescents with their levels on comparable tests of achievement was used as a preliminary measure of the validity of this instrument.

Not only have developmentally disabled persons generally been shown to hold jobs requiring less skill (Richardson, 1978), but specific training is often necessary for them to reach their potential in both work and leisure situations (Clark, 1980; Marion, 1979). The systematic assessment of work skills and work-related social competence has therefore become a matter of serious concern, particularly given recent emphasis on vocational education of the handicapped under Public Law 94-142 (Bellamy, Wilson, Adler, & Clark, 1980).

Recent approaches have tended to emphasize assessment of a variety of cognitive and social skills deemed essential to vocational success. Quinones (1978) has examined the use of a standard psychological battery of tests and found that achievement test measures were among the best predictors. Schreiner (1978) also found that measures of cognitive ability were related to work performance but that samples of behavior under actual work conditions were among the best predictors of work success, and task analysis has been favored as an effective means of vocational assessment (Friedenberg & Martin, 1977). Personal independence has likewise been cited as a good predictor (Cunningham & Presnall, 1978), as have verbal manners and communication skills, especially when these are considered in relation to the appropriateness of a developmentally disabled person's job assignment (Malgady, Barcher, Towner, & Davis, 1979). Exemplary programs of assessing a handicapped person's work potential by observing his or her response to actual job training have also been developed (Bellamy, Horner, & Inmen, 1979; Gold, 1973). These appear to have the advantage of more immediate generalization of results to actual work situations.

Much in these findings points to the wisdom of assessing a developmentally disabled person's vocational potential by evaluating his or her academic and social skills within the context of a work-related situation. Conversely, there also seems much to suggest that instruction of handicapped students, whether in regular or special classrooms, should be increasingly carried out with a view towards *application* of their academic skills. The generalization of skills learned in a classroom setting may then be less problematic when these skills eventually have to

BEST COPY AVAILABLE

be applied to actual work or leisure situations (Halpern, 1979; Porter, 1980; Stodden, Casale, and Schwartz, 1977).

The present study is a report on the development of an instrument to assess *applied* academic and social skills, plus an analysis of data on the concurrent validity of the instrument used for this technique.

The instrument, hereinafter referred to as the "Applied Assessment Instrument" or AAI, was developed within the context of a prevocational shop classroom for developmentally disabled adolescents hospitalized for psychiatric disorders and has been used for the past two years to evaluate the readiness of these students for various prevocational or hobby activities.[1] The AAI was not only intended to be applicable to a variety of prevocational situations for developmentally disabled adolescents and young adults but also to serve as a model for *feedback* to classroom teachers on handicapped students' ability to apply the academic and social skills learned in regular or special classroom settings. Results are therefore used not only to assess the student's academic and social readiness to function in various shop activities but, just as importantly, to provide data to the student's concurrent classroom teachers on any prevocational academic competencies which the student does not yet seem to have mastered.

AAI is designed to assess reading, number, social, and motor skills using materials taken from typical work or hobby projects. These will be discussed in more detail. The present study includes a comparison of the scores of developmentally disabled adolescents on five areas of the AAI (applied reading, comprehension, direction following, number skills, and social skills) with their subtest scores on the California Achievement Test (Tiegs & Clark, 1963) in reading vocabulary, comprehension, arithmetic reasoning, and arithmetic fundamentals and with a behavior rating by their classroom teachers.

Methods

Subjects

Subjects for the study were 30 developmentally disabled adolescents (17 male and 13 female) admitted for evaluation and short-term treatment on one inpatient ward of the Mental Retardation and Child Psychiatry Program, UCLA Neuropsychiatric Institute (NPI). Mean age of the subjects was 14.5 years (range 12.2 to 20.8 years). All subjects having complete test records were selected from a total population of 48 patients admitted to the ward over a 2-year period from 1977 to 1979. All subjects had developmental disabilities (primarily mild retardation and/or serious emotional disorders), and their IQ range was approximately 50 to 80.

Setting and Testing Procedures

All subjects were enrolled in NPI School classrooms for 2 to 3 hours each day, and one of their 2-hour periods each week was spent in the prevocational shop program of the NPI School. All subjects were enrolled in the shop program for a minimum of 1 month with a range from 1 to 5 months. A more complete description of the total school program is provided in Forness (1977). The administration of each subject's AAI was routinely done in an individual session by the shop teacher during the first 2 weeks of each subject's admission. California Achievement Tests (CAT), at the Upper Primary or Elementary levels, were administered in individual sessions by the student's classroom teacher during the same 2-week period. A behavior rating was also done by two classroom teachers who rated the same student independently within the first 2 weeks of admission to the classroom.

Grade level equivalents for each CAT subtest were used in the data analysis. Each student's classroom teachers (in his or her academic classroom, not the shop program)

---

[1] The instrument developed for this technique is the "Assessment of Applied Academic and Social Skills" of the UCLA NPI (Neuropsychiatric Institute) School Prevocational Shop Classroom. The title has been shortened for convenience herein. Sample copies of the instrument and accompanying answer and summary report sheets, as used in the NPI School, are available without charge and may be obtained by writing the authors.

BEST COPY AVAILABLE

were asked to rate the student on a scale routinely used by the NPI School staff to measure the student's behavior and peer relationships in the academic classroom. There were eight items, each rated on a 7-point Leikert scale. The eight items were task attention, independent working, ability to communicate needs effectively in the classroom, ease of classroom management, participation in group classroom activities, age-appropriate peer interactions, popularity with classroom peers, and ability to develop reciprocal friendships. As indicated above, two teachers (or one teacher and an assistant teacher) rated the student independently on these scales, and results from both teachers were averaged for each scale. Possible scores ranged from 8 (indicating poor adjustment) to 56 (indicating good adjustment).

Assessment of Applied Skills

The techniques used on the AAI to measure applied academic and social skills were based on a number of years observation and experience in the NPI prevocational shop classroom. They were also based on information about the demand characteristics of vocational and prevocational settings from which adolescent patients had been referred and in which they were subsequently placed after discharge from the NPI program. Several versions of the Applied Assessment Instrument were piloted before the present version was developed over 2 years ago.

The first section of the AAI is designed to assess three skills related to reading: *applied reading, comprehension,* and *direction following.* "Reading" skills are measured by having the student read aloud three short paragraphs of directions from actual projects at three levels of difficulty: primary grades (from a project on building a wooden toy truck), upper elementary (doing simple needlepoint), and junior high (constructing a plywood bookrack). "Comprehension" is measured by having the student repeat these directions in his or her own words, and "direction following" is measured by having the student actually carry out these directions using the actual materials described. Operational definitions of specific data-based performance were converted to 6-point Leikert-type scales, for each ability on each level of paragraph difficulty. Possible scores for the three reading-related abilities (applied reading, comprehension, and direction following) thus ranged from 3 to 18 points each.

The second section is designed to assess *applied number skills* related to counting, numbering, measurement, and telling time. Applied number skills are measured by counting pieces or materials depicted on a sheet of directions, pointing to numbered items depicted on drawings, telling time, measuring items with a ruler, and identifying simple fractions and decimals relating to measurement. Assessments are obtained in ten areas, each measured on a 5-point Leikert scale, such that a possible score ranges from 10 to 50 points for this section. The third section, *social skills,* contains six areas which are assessed by observing the student's concentration, frustration point, ability to accept guidance, ability to work independently, care and respect for tools, and cooperation while being tested. Each area is assessed on a 5-point scale with a possible score ranging from 6 to 30 points.

The fourth section is devoted to *applied motor skills* which are measured by observing the student's ability to use common tools and rating his or her performance on structured tasks. Skills observed are the student's use of a screw driver to attach different sized screws to wood, threading nuts and bolts with a wrench, cutting a board at differing angles with a handsaw, and driving various nails into wood with a hammer. Skill is also observed in threading different sizes of needles, drawing a simple grid, and the use of scissors to cut simple patterns out of both cloth and paper. This motor skills section, however, was not included in the present analysis which focuses primarily on academic and social skills. It should be noted that total assessment time on the AAI for each individual was approximately 45 minutes.

Results

Table 1 presents the mean, range, and standard deviations of the CAT scores, classroom teacher ratings, and applied academic and social scores for the entire sample. Note

BEST COPY AVAILABLE

TABLE 1

Means, Ranges, and Standard Deviations for CAT Subtests, Teacher Ratings, and Applied Academic and Social Assessments

Classroom Measures	*Mean*	*Range*	*SD*
CAT reading vocabulary	4.60	(1.7 to 9.0)	1.88
CAT reading comprehension	4.55	(1.8 to 7.9)	1.44
CAT arithmetic reasoning	4.52	(1.2 to 7.9)	1.38
CAT arithmetic fundamentals	4.37	(.9 to 8.7)	1.65
Classroom teacher's behavior rating	16.1	(10.0 to 24.5)	4.14
NPI applied academic and social assessment			
Reading	14.9	(3 to 18)	3.23
Comprehension	11.8	(3 to 18)	4.91
Direction following	10.6	(3 to 18)	4.41
Number skills	29.6	(24 to 49)	7.31
Social skills	22.9	(9 to 30)	6.59

that the mean achievement scores of the subjects are at the fourth grade level and that there appears to be considerable variability in both the classroom teacher ratings and the applied academic and social scores.

Spearman rank-order correlations (*rho*'s) were computed between CAT subtest scores and corresponding academic areas on the applied assessment instrument, and these are presented in Table 2 along with correlations between ratings of classroom teachers and applied social skills. Note that generally AAI reading scores do not correlate as highly with CAT reading vocabulary as they do with CAT reading comprehension nor do AAI number skills correlate as highly with CAT arithmetic fundamentals as they do with arithmetic reasoning.

Discussion

The moderate but generally significant correlations found between classroom measures and assessment of similar academic and social skills in an applied situation appear to suggest at least some support for the concurrent validity of the AAI as described herein. Correlations of a much higher order might, in fact, have been interpreted as *not* supporting the use of the instrument since a more complete overlap of the two types of measurement would suggest that the more

TABLE 2

Correlation Coefficients (*rho*'s) Between Corresponding Classroom and Applied Assessment Measures

	Applied Academic and Social Assessment				
Classroom Measures	*Reading*	*Comprehension*	*Direction following*	*Numbers*	*Social skills*
CAT reading vocabulary	.345*	.449**	.310*		
CAT reading comprehension	.527**	.638**	.499**		
CAT arithmetic fundamentals				.363*	
CAT arithmetic reasoning				.514**	
Classroom behavior rating					.560**

*Significant at the .05 level
**Significant at the .01 level

BEST COPY AVAILABLE

traditional use of achievement scores and teacher ratings are sufficient to predict prevocational or vocational success. It is of interest in this regard that the basic skills of reading vocabulary and arithmetic fundamentals did not seem to relate as highly to AAI skills as did reading comprehension and arithmetic reasoning, which tend to measure the *application* of basic academic skills.

In the same vein, direction following, as measured on the AAI, did not tend to correlate as highly with any CAT reading subtest as did the AAI measures of reading and comprehension. These findings may suggest, at least in this instance, a possible continuum ranging from acquisition of basic academic skills (e.g., reading recognition) through their application in an academic context (e.g., reading comprehension) to their ultimate application in a practical context (e.g., following written directions in a prevocational situation).

The fact that at least moderate relationships were found using a relatively heterogeneous adolescent population with relatively variable behavior seems also to be significant. The assessment of academic and social skills under work-related conditions would seem particularly essential to placement decisions involving developmentally disabled persons with obvious problems in deportment. The AAI, or instruments similar to it, might be useful with just these types of individuals since their ability both to profit from academic instruction as well as to apply what skills they do learn may be more limited than developmentally disabled students with few behavior problems (Szymanski & Tanguay, 1980).

It should be stressed, however, that findings illustrating the validity of the AAI or even the instrument itself are not as important as the concept of assessing a handicapped student's abilities to *apply* academic and social skills. Any assessment of this type is essential for two purposes. First, it enables a shop or vocational instructor to ascertain whether a student has mastered the academic or social competencies essential to either prevocational classroom situations or on-the-job training. Secondly, it provides feedback to classroom teachers on whether their instruction is actually *relevant* to the student's current career or vocational needs. In both instances, there are clear implications for the instruction of developmentally disabled or other types of handicapped students, either in regular or special class settings. Particularly as such students reach junior high school ages, or even prior to that point, teachers need to be concerned more directly with the generalization and application of skills learned in their classrooms. Instruction or remediation in such skills may be more effective if begun concurrently with, or even prior to, the student's prevocational or vocational training. Assessment of prevocational academic skills thus becomes critical to the task of providing the student's teacher with feedback as a basis for developing more relevant individual instructional programs in the classroom.

References

Bellamy, G. T., Horner, R., & Inman, D. *Vocational habilitation of severely retarded adults: A direct service technology*. Baltimore: University Park Press, 1979.

Bellamy, G. T., Wilson, D., Adler, E., & Clark, J. A strategy for programming vocational skills for severely handicapped youth. *Exceptional Education Quarterly*, 1980, *1*, 85-98.

Clark, G. M. Career preparation for handicapped adolescents: A matter of appropriate education. *Exceptional Education Quarterly*, 1980, *1*, 11–18.

Cunningham, T., & Presnall, D. Relationship between dimensions of adaptive behavior and sheltered workshop productivity. *American Journal of Mental Deficiency*, 1978, *82*, 386–393.

Forness, S. A transitional model for placement of handicapped children in regular and special classes. *Contemporary Educational Psychology*, 1977, *2*, 37–49.

Friedenberg, P., & Martin, A. Prevocational training of the severely retarded using task analysis. *Mental Retardation*, 1977, *15*, 16-20.

Gold, M. W. Research of the vocational habilitation of the retarded: The present, the future. In N. Ellis (Ed.), *International Review of Research in Mental Retardation* (Volume 6). New York: Academic Press, 1973.

Halpern, A. S. Adolescents and young adults. *Exceptional Children*, 1979, *45*, 518–523.

Malgady, R., Barcher, P., Towner, G., & Davis, J. Language factors in vocational evaluation of mentally retarded workers. *American Journal of Mental Deficiency*, 1979, *83*, 432–438.

BEST COPY AVAILABLE

Marion, R. Leisure time activities for trainable mentally retarded adolescents. *Teaching Exceptional Children*, 1979, *11*, 158–160.

Porter, M. Effect of vocational instruction on academic achievement. *Exceptional Children*, 1980, *46*, 463–464.

Quinones, W. A test battery for assessing the vocational competency of moderately mentally retarded persons. *Mental Retardation*, 1978, *16*, 412–415.

Richardson, S. Careers of mentally retarded young persons: Services, jobs, and interpersonal relations. *American Journal of Mental Deficiency*, 1978, *82*, 349–358.

Schreiner, J. Prediction of retarded adults' work performance through components of general ability. *American Journal of Mental Deficiency*, 1978, *83*, 77–79.

Stodden, R. A., Casale, J., & Schwartz, S. Work evaluation and the mentally retarded: Review and recommendations. *Mental Retardation*, 1977, *15*, 25–27.

Szymanski, L. S., & Tanguay, P. (Eds.) *Emotional disorders of mentally retarded persons: Assessment, treatment, and consultation*. Baltimore: University Park Press, 1980.

Tiegs, E., & Clark, W. *California Achievement Tests*. Monterey, CA: California Test Bureau, 1963.

STEVEN R. FORNESS *is Professor and Special Education Director, UCLA Neuropsychiatric Institute;* **ROBERT L. THORNTON** *is a Demonstration Teacher, and* **AMY A. HORTON** *is an Assistant Teacher, Neuropsychiatric Institute School, University of California, Los Angeles.*

Handbook for Disseminators

The 1980 *Guide To Packaging Your Educational Program*, a 91-page handbook for disseminators, is focused on the processes for project staff of a validated exemplary program to follow when preparing for state or national diffusion.

Adapted from a 1976 publication, "Packaging Your Educational Program," the handbook has been improved and updated. It covers content, how to plan staffing, time, and budget; production of print and audiovisual materials; and problems of distribution including copyright, releases, pricing, inventory, shipping, and marketing. An appendix describes a complete inventory of materials for a comprehensive dissemination package.

Copies are available from the Far West Laboratory's Order Department, 1855 Folsom Street, San Francisco, CA 94103, for $5.50 prepaid.

BEST COPY AVAILABLE

DIAGNOSING INSTRUCTION

Siegfried Engelmann, B.A.
Alex Granzin, Ph.D.
Herbert Severson, Ph.D.
Oregon Research Institute, Eugene

Traditional diagnostic approaches usually occur outside the instructional context and focus on the learner. Accurate conclusions about the learner, however, can only be reached after an adequate diagnosis of instruction. Failure to consider instruction as a variable results in diagnoses that lack specific implications for teaching. Assumptions and procedures critical to an adequate diagnosis of instruction are discussed and applied to a variety of examples. The process of derivation of remedies and its relationship to the diagnostic process is illustrated with examples. Implications for the diagnostician and the consequences of failing to provide adequate instructional diagnoses are discussed.

VARIABLES OF DIAGNOSIS

Traditional procedures used to diagnose the learner are extremely weak because they do not assess the instruction the learner receives; they only assess the learner. Typically, a diagnosis is called for when there is trouble, generally because whatever the teacher is doing is not working. The learner is not "growing," "developing," "interacting," or "behaving" in an appropriate manner. A diagnosis is achieved by removing such learners from the instructional setting and giving them a series of tests that provide a sample of the learner's behavior. The assumption is that the sample provided by tests is somehow better than other samples. While the interpretation of the sample ostensibly tells something about the learner, the diagnostician never draws conclusions about persons or factors other than the child, e.g., "The child's performance on the Bender Gestalt clearly indicates that the teacher is very poor at classroom management." On the contrary, interpretation always tells about the learners — their predisposition, mental abilities, skills, personality, intelligence, sensorimotor performance, and so on.

There is no severe problem with the diagnosis until a remedy is drawn from it. For while the sample of behavior may be useful for classifying the learner, it fails to suggest adequate remedial action. Every special educator who has played the remedy game recognizes the problems encountered at this juncture. "It seems," the psychologist says thoughtfully to the teacher, "that you should work on auditory sequencing."

"Like what?"

"Well, counting."

"We do it."

0022-4669/79/1304-0003$01.00/0

BEST COPY AVAILABLE

"Well, what about completing words in a sentence?"

"Like what kind of sentence?"

" 'We find cows in the . . . .' "

"He can't do it."

"That's probably what you have to work on."

"How?"

At this point, diagnosticians are so far beyond the information provided by the test, and so far from the realities of the teaching situation, that if they're smart they'll retreat from the concrete details that teachers need to the more comfortable generalities that characterize staffings. Actually, the diagnostic procedure is a perfect charade if the goal of diagnosis is to lead to a remedy. For a remedy to follow from the diagnosis, the remedy must provide teachers with information that they don't have. It must tell them what, if anything, they can do to improve learner performance. It should not tell them what they already know, that the learner has some sort of deficiency. Teachers know precisely how the deficiency manifests itself, what the learner has trouble learning, and how the learner responds to different situations. Teachers, however, are not usually insulted by a diagnosis that may tell them far less than they already know if that diagnosis assures them that (a) the instructions they have provided are adequate and (b) the learner's failure is the result of some basic flaw in the learner. Traditional diagnoses are designed to provide these assurances. Unfortunately, they carry very limited remedial implications because they do not specify (a) the extent to which the learner's failure is caused by poor instruction and (b) precisely what the teachers could do to remedy an observed problem.

The traditional diagnosis is incapable of expressing remedies in the basic units the teacher manipulates when teaching. The teacher achieves teaching (or changing behavior) *not* by manipulating neurons, the learner's past history, or internal processes of any sort. Rather, the teacher achieves behavioral change only by manipulating environmental events. This point is extremely important. It follows that the remedy must clearly imply manipulation of those environmental events. It must tell teachers what they are doing wrong and what types of different teaching behaviors they should implement. The remedy must be specific and concrete because teaching always involves specific, concrete acts. The instruction "Teach seriation" does not tell the teacher what to do, when to do it, how to respond to specific errors, how to sequence examples, how to review, or how to reinforce. Yet, in order to "teach seriation," the teacher has to present examples in sequence, say specific words, and respond in some way to the learner's attempts.

Traditional diagnosis assumes that all relevant information comes from a study of the learner. An equally tenable position is that all relevant information comes from a study of the instruction the learner receives, not from the learner. Neither position is reasonable. The learner's behavior is influenced by two major factors: (a) the innate capacity or predisposition of the learner and (b) the instruction the learner receives. In other words:

$$B \text{ (Behavior)} = P \text{ (Predisposition)} + I \text{ (Instruction)}.$$

BEST COPY AVAILABLE

In this formula, *Predisposition* would include all factors that remain unaffected by instruction. (If the learner exhibits the same problem regardless of instructional approach, the problem is controlled by predisposition.) *Instruction* refers to the effects of teaching — intentional or unintentional. (If the problem can be eliminated through instruction, the problem is controlled by instruction.)

When we observe a given behavior, we cannot specify the extent to which it is controlled by predisposition or by instruction. We may assert that it is controlled primarily by one factor or another, but our assertion is based on ignorance, not on fact. We might conclude that the learner has a perceptual problem when in fact the observed behavior has been caused by poor instruction and can be eliminated through instruction.

One way to extricate ourselves from this diagnostic dilemma would be to rule out one of the two variables (P or I). If we could eliminate the learner's predisposition from the observed behavior, we could use the formula: $I = B - P$. If we could rule out the instructional influence and look only at the learner's predisposition, we could use the formula: $P = B - I$. Unfortunately, we cannot remove the learners from the effects of the instruction they receive or from their innate predispositions. We must therefore try to control the variables P and I in some other way, and the control must be designed so that it doesn't require removing a variable (or pretending that we remove it).

There is no obvious way to control P by improving it, maximizing it, or redesigning it so that we know precisely how it works. We can, however, control I. To do this, we maximize I, which means that we design it so that it works across a wide range of learners. When we design I so that it is basically faultless and incapable of contributing to the learner's behavior problem, we can draw conclusions about both P and I. If B changes greatly with I controlled, we conclude that I is the primary factor in determining B. If B does not change greatly, we conclude that P is primarily responsible for the status of B.

We have now gone full circle and have come to the central diagnostic problem. We wish to control I. First, we must determine what type of control is needed and to what extent I is in need of maximization. In other words, we must diagnose instruction. The diagnosis perforce must be of instruction, not of the learner. The instructional diagnosis involves two steps: (a) interpreting instruction the learner receives according to the minimum-knowledge assumption, and (b) testing the minimum-knowledge assumption by providing a maximum-knowledge test.

The minimum-knowledge assumption

When we view instruction, we see the learner responding to different things the teacher does. The teacher presents tasks, and the learner responds. The minimum-knowledge assumption holds that the learner uses the least possible knowledge required to produce the various behaviors we observe. For example, the teacher says, "Open the door, Henry," and points to the door. If Henry opens the door, we cannot assume that Henry understands the command, "Open the door," because Henry was not required to respond only to the

BEST COPY AVAILABLE

words. He could have responded only to the teacher's pointing. The minimum-knowledge assumption holds that Henry did just that — responded to the point, not to the words.

The maximum-knowledge test

The minimum-knowledge assumption is perfectly consistent with the observed behavior. Further, it carries implications for instructional control. That is, we can design a test that eliminates the teacher's point and presents only words. This test requires knowledge of the words — the maximum-knowledge test. The minimum knowledge assumption identifies the simplest mechanism that accounts for what the learner does. The test of maximum-knowledge restructures the situation so that the learner cannot use the simplest mechanism.

Failure to make the minimum-knowledge assumption results in failure to test the possibility that the learner is responding to messages other than the ones the teacher intends. Let's say the teacher hands a learner a piece of candy and says, "Eat this." The traditional interpretation assumes that if the learner ate the candy, the learner understood the directions (maximum-knowledge). The minimum-knowledge assumption holds that the learner is responding to the candy in the hand, not to the words; in other words, the learner would eat it no matter what the teacher said. Then the minimum-knowledge assumption can be tested (perhaps by handing the learner candy and telling him/her, "Don't eat it," or "Shut the door"). The minimum-knowledge assumption implies the maximum-knowledge test. If the minimum-knowledge assumption is not made, however, the role of the instruction in the learner's performance is not tested.

Two points about the diagnosis of instruction should be noted:

1. The minimum-knowledge assumption is just that — an assumption, not a fact. The assumption should be tested, and the test will clearly determine the extent to which it is confirmed.

2. The minimum-knowledge assumption can be made about written descriptions, but the use of these descriptions in formulating diagnoses is not efficient because they do not provide sufficient detail about what happens during instruction. Written descriptions only tell about those details of which the writer is aware. If the writer is unaware of such details as sequencing items, the written instruction obviously will not mention these. Also, written descriptions are often useless because teachers do not follow them.

APPLICATIONS

Below are three applications of the diagnostic procedure, relatively elementary examples. The same basic procedure can be used in more complex situations, however. The first two examples involve situations in which a problem has been identified. Situation 1 involves a learner whose behavior is not changing in the expected way. Situation 2 concerns a learner who seems to have trouble following instructions. Situation 3, however, involves a learner who is performing acceptably but who exhibits a problem when the procedure used in the previous examples is applied. The minimum-knowledge assumption iden-

BEST COPY AVAILABLE

tifies possible misinterpretations that are conveyed through instruction. The maximum-knowledge test confirms that the deficiency in instructional procedures has resulted in a deficiency in the learner's understanding. The point is that the diagnostic procedures can be applied before a problem is identified and that preventive measures can be designed to obviate the instructional deficiency.

Situation 1

The teacher is trying to teach or reinforce number skills of low-language children by playing a game with oversized dominoes. After placing a domino in the middle of the table, the teacher says the number shown by his/her domino. If a child has a domino with a matching number of dots, the child is to place that domino in the middle of the table.

The problem child in the group throws out a domino every time the teacher places one on the table. The teacher's response is simply to ignore the inappropriate responses by pushing the domino back to the child and continuing the task. When the child's response is appropriate, the teacher reinforces it. The teacher's attempts to "shape" the child's behavior have not worked.

Minimum-knowledge assumption. To formulate the assumption we ask, "What is the least amount of knowledge learners could possess and perform in the observed way?" They would not have to count the dots on the dominoes and would not have to understand that they are supposed to throw out a domino only if it matches. They could perform in the observed manner if they operated from the principle: "Throw out your domino each time the teacher presents one. From time to time, you'll be reinforced."

We assume, therefore, that the learner operates from this behavioral rule and that the learner has no knowledge of numbers, matching, or the rules of the game.

Maximum-knowledge test. A number of maximum-knowledge tests are possible here. The requirement for each is that it must test the child's knowledge of whether the number of dots match. For instance, we could provide the learner with a row of dominoes and hand him/her one domino, which the child must place next to the domino with the same number of dots.

Situation 2

Jenny performs successfully when the following tasks are presented in random order: "Touch your head . . . touch your nose . . . stand up . . . pick up the fork . . . pick up the pencil."

Jenny has difficulty with tasks such as, "Put the pencil under the chair," and "Touch the pencil." (She puts the object handed to her *on* the chair, and she picks up the pencil.)

Minimum-knowledge assumption. Jenny peforms by cueing on the last word in the command. She cannot perform on preposition tasks because words other than the last word provide instruction about what to do. She cannot perform on "Touch the fork" because she ignores the word *touch*.

Maximum-knowledge test. The maximum-knowledge test must first deter-

BEST COPY AVAILABLE

mine whether Jenny could follow two different commands involving the same object. Could Jenny, for instance, "touch the fork" as well as "pick up the fork?" The test would involve random trials of these commands. If the minimum-knowledge assumption is correct, Jenny would pick up the fork on every trial, even after the correct response to *touch* had been modeled and emphasized in the commands.

Situation 3

Ann, a junior-high student, performs well on a variety of arithmetic tasks. She works simple algebra problems; however, the program that she uses requires application of laws, such as associative and distributive, and always requires problems to be solved for X or A (not for $2X$ or $\frac{1}{2}A$).

Minimum-knowledge assumption. Ann has learned strategies for solving problems that would not permit her to solve a problem such as: $4X = 7$; $3X =$ __ . This minimum-knowledge assumption is based solely on the program, not on an observed behavioral deficiency. The assumption is that if Ann is "normal," it is quite probable that she developed a strategy that would not work well in solving for values other than X.

Maximum-knowledge test. The most direct test would be to require Ann to solve the same problem for different values:

$$\begin{array}{llll} 4X = 7 & 4X = 7 & 4X = 7 & 4X = 7 \\ 3X = \_\_ & {}^{2}/_{3}X = \_\_ & 7X = \_\_ & {}^{1}/_{2}X = \_\_ \end{array}$$

Remedies

Remedies follow logically from learners' performance on the maximum-knowledge test. If learners fail the test, they must work on tasks or activities that can account for passing the test.

If the domino player did not pass the test that required matching a domino with the appropriate one in the teacher's display, the task would be simplified (reducing the number of dominoes displayed by the teacher), and the preskill of counting the dots would be taught and then applied to the matching procedure. Initially, the structure of the task would involve steps such as: "Count the dots on your domino . . . How many dots on your domino? . . . Show me the other domino that has four dots . . . Put your domino next to the domino that has four dots . . . Good job." Later, the steps would be "faded" and the learner would be required to perform without prompting. Once he/she had demonstrated proficiency required to pass the maximum-knowledge test (in which he/she would be given different dominoes to be placed next to those that match), the original domino game could be reintroduced, perhaps with new rules: "If you put out a domino that does not match, I get one of your points."

Again, if Jenny failed the maximum-knowledge test, the implied remedy is to teach the skills required to pass the test. To achieve this, we might first teach her single word commands, such as "touch." We would initially present unfamiliar objects or those that are not easily picked up. The teacher would say the word "touch" and then demonstrate or model the response. Jenny would next be tested. After performing acceptably on these objects, objects such as

BEST COPY AVAILABLE

forks and candy would be introduced. Next, she would be taught to discriminate between the command "touch" and the command "pick up." Both commands would be given for the same object. For instance, a fork would be presented, and Jenny would be given these commands: "Touch . . . touch . . . pick up . . . touch . . . pick up . . . pick up . . . pick up . . . touch" etc. After Jenny performed acceptably on this type of task, tasks that name both the object and the action would be introduced: "Touch the ball . . . Pick up the book . . . Touch the fork . . . Touch the candy." When Jenny can perform these tasks, she has passed the maximum-knowledge test. To respond correctly, she must attend to and understand the meaning of the commands, which means that she must have maximum knowledge.

The remedy for the poor arithmetic sequence would be to teach Ann a strategy that would permit her to solve problems that require solutions for values other than *X*. A possible strategy is to show Ann that a problem of this type:

$$\begin{aligned} 4X &= 7 \\ 3X &= \_\_ \end{aligned}$$

is simply a ratio problem with equivalent fractions on either side of the equal sign:

$$\frac{4}{3} = \frac{7}{\_\_} .$$

The game involved in solving the problem is to find the number that 4 must be multiplied by to change it into 7. We must multiply 3 by the same value to get the answer.

Once Ann has applied this analysis to various problems (both ratios and those involving letters), she should have no trouble with the maximum-knowledge test.

ASSUMPTIONS, TESTS, AND REMEDIES

The procedures outlined show the relationship between the minimum-knowledge assumption, the maximum-knowledge test, and the remedy. The examples of remedies are based on situations in which the learner fails the maximum-knowledge test. If the learner passes the test, of course, no remedy is implied; we simply conclude that our assumption of minimum-knowledge was not confirmed by the test. Possibly, the test will provide only a partial confirmation of the assumption, which means that the learner will fail only part of it or only some types of items. Performance will imply what we must teach before the learner performs adequately on the missed part or items.

DIAGNOSIS AND REMEDY

The purpose of instructional diagnosis is to determine aspects of instruction that are inadequate, to find out precisely how they are inadequate, and to imply what must be done to correct their inadequacy. The assumption of minimum-knowledge is central to the diagnosis. If it is not made, the instruction is auto-

BEST COPY AVAILABLE

matically exonorated, and the fault for poor performance is automatically placed on the learner. If we fail to recognize that instruction permits and reinforces domino players for throwing out their domino on every trial, we may conclude that they are slow learners, that they do not respond to the "punishment" of teachers ignoring their inappropriate responses. The teacher's procedures, although faulty, become exonorated, and the diagnosis shifts to questions of why the learner tends to respond in such a strange way.

By failing to make the minimum-knowledge assumption about Jenny's performance, we are left with the uncomfortable conclusion that she is apparently capable of responding to commands; however, this is not always so. Unless we note that her performance is perfectly consistent with what she has been reinforced for doing (responding to commands in which the last word conveys all the important information), we must try to account for her erratic behavior. Our attention moves from the instruction to Jenny's personality, learning style, developmental pattern, or whatever.

Similarly, failure to apply the minimum-knowledge assumption to Ann prevents us from further tests of the program's adequacy. Instead, when she develops serious problems in arithmetic, we conclude that her problems are caused by an internal mechanism, not by the program: "Ann lacks aptitude."

In every case, our conclusions are premature and unfounded. They represent a *possible* interpretation; the other, and equally possible, interpretation is that the learner responds in a way consistent with the instruction presented.

When instructional deficiencies are not identified, the teacher may understandably become frustrated. Even though the teacher works harder on following directions, Jenny persists in making the same mistakes. Playing the domino game more frequently does not seem to improve the learner's performance. Reviewing simpler problems does not seem to help Ann work the more difficult ones.

In short, the teacher draws dangerous conclusions about the learners; instruction apparently does not seem to change their responses; therefore, the problem must be with their predispositions.

DIAGNOSING THE LEARNER

The only valid way to draw conclusions about deficiencies involves first determining the degree to which the learner's performance is controlled by instruction. We must go through the following steps:

1. Diagnose the instruction the learner receives (observing instruction using the minimum-knowledge assumption and then providing the maximum-knowledge test).
2. Provide instructional remedies that account for the skills required by the maximum-knowledge test.
3. Observe discrepancies between the learner's performance and that of a normal learner.

The first two steps factor out instruction and identify the extent to which it is responsible for the observed behavior. We can then indicate the extent to which performance is caused by the learner's basic makeup — the third step. If the

BEST COPY AVAILABLE

learner's performance does not improve in response to instruction designed to teach maximum-knowledge skills, we can conclude that the performance deficiency is caused primarily by predisposition. If the learner learns in a so-called normal way, however, we can conclude that predisposition plays a minor role.

We must temper our conclusions with an understanding of learning. For example, when a remedy requires the learner to abandon behaviors that have been reinforced for years, we would expect learning to proceed slowly, even if the learner is normal. Relearning requires much more practice than initial learning. Moreover, learning unfamiliar discriminations requires considerably more practice than does learning familiar discriminations (Engelmann & Granzin, in press). In the case of a learner who has never attended to verbal utterances, we can assume that initial learning (even for a normal learner) would be quite slow (requiring hundreds of trials to master only a few discriminations). When our goal is to assess normality, we can use any deficiency as the basis for concluding that the learner is below normal. If we wish to find out the extent to which this behavior implies a persistent and irreversible deficiency, we must first *teach* the learner and observe responses.

IMPLICATIONS FOR THE DIAGNOSTICIAN

Traditionally, the diagnostician's role has been to provide the teacher with a fine-grained classification of the learner's problems. As noted, the instruments focus solely on the learner in an attempt to tell what type of deficiency he/she has, to relate it to the norm, and to convert the teacher's casual observations of the problem into more scientific language.

The diagnostician must recognize that this orientation is faulty. The purpose of diagnosis is not primarily to provide a refined statement of what the teacher already knows, nor is it to sanction the idea that merely because learners possess a deficiency, they are automatically the cause of the deficiency. The diagnostician can perform a useful role only by recognizing that the goal of instructional diagnosis (as opposed to a demographic statement about the child's deficiency) is to imply an instructional remedy.

Accepting this orientation is to accept the fact that *teachers need help*. Not all problem children are abnormal. Not all the procedures teachers use to correct these children are well conceived. And certainly not all the instructional sequences they use are faultless — or even adequate. This being so, the remedial focus must move to the teachers. For the children to receive effective remedies, their teachers must behave in different ways. Teachers probably will not change their behavior, however, unless they are taught to do so. If they are not taught, new behavior problems and learning problems among children may be created at an expected rate.

Reference

Engelmann, S., & Granzin, A. Unfamiliar learning. In J. Pickett & A. Risberg (Eds.), *Proceedings of the Gallaudet Conference on Speech Processing Aids for the Deaf*, in press.

BEST COPY AVAILABLE

The New Direction in Educational Research: Alterable Variables

by Benjamin S. Bloom

One of America's leading educational researchers discusses certain alterable variables that have brought about major changes in our views of learners and their amazing potential for learning. He believes the list of alterable variables will be rapidly expanded in the Eighties, bringing profound changes in school and society.

A major revolution has taken place during the past decade in educational research and our understanding of some of the factors that directly influence learning in or out of the schools. As a result, student learning can now be improved greatly, and it is possible to describe the favorable learning conditions that can enable virtually all students to learn to a high standard. Researchers who were at one time concerned about providing equality of educational opportunity for students now speak of the learning conditions that can bring about equality of educational outcomes for students. And such educational outcomes are at very high levels of attainment.

The direct application of this new understanding to teachers and the schools has taken place very quickly in some places, but in others it may await the presence of several conditions. Application of the new understanding will depend upon the leadership in the schools, the need for improvement in learning, and the role to be played by schools of education and teachers of teachers. In some countries the leadership in applying the new research to the schools has been assumed by curriculum centers that weave these new ideas into the instructional material and into the instructional processes for the new curriculum. This procedure has been especially effective in curriculum centers that provide inservice education to teachers for the new curriculum.

At least four methodological features account for the striking qualities of these new research developments. The simplest of these is the movement from a study of the characteristics of teachers and students to direct observation of learning taking place in the interactions between teachers and students in the classroom. Perhaps, to put it in the most direct terms, it is a movement from the study of the actors (teachers and students) to the study of teaching and learning as they take place under specific environmental conditions.

Increasingly, educational researchers are performing experimental studies under classroom conditions in which selected variables are studied in terms of the processes involved as well as the changes they produce in both teachers and learners. Central to these studies is the concern about the causal links between the process variables and the qualitative and quantitative changes in the learning of students. These are very different from the pre-post demonstration studies of an earlier period, since they center on the teaching and learning processes that take place between the pre and post measurements.

Third, these experimental studies are guided by models and theories that embody causal links. These models and theories are the basis for the hypotheses to be tested and the research designs to be used. Methods of strong inference are used to relate theory and models to classroom observations and experiments in such a way that a series of studies under varying conditions is necessary to test and improve the theory and the models.

But perhaps the most important methodological change is the movement from what I have termed stable or static variables to variables that are alterable either before the teaching and learning processes or as a part of these processes. I consider this shift in the variables used as central to the new view of education. This shift enables researchers to move from an emphasis on prediction and classification to a concern for causality and the relations between means and ends in teaching and learning. This new concern has resulted in new ways of understanding, explaining, and altering human learning. The search for alterable variables, and the causal processes by which they can be altered, is a relatively recent step in educational research. I am confident it will be central in educational research of the next decade.

In the following pages I shall describe some of these new alterable variables, contrast them with the nonalterable variables they replace, and indicate the gains in education to be secured from this new approach to educational research and educational practice.

Available Time vs. Time-on-Task

We have always recognized time as a central factor in all learning. Schools have allocated a certain number of years for different subjects such as reading, literature, arithmetic, science, or social studies. In addition, the schools determine the number of school days in each school year and the numbers of hours per day or week that will be assigned to each part of the curriculum. Time in the sense of years, days, and hours available for school learning becomes a relatively fixed or stable variable. To make significant alterations in these time allocations requires major legal, economic, and other policy changes at the state or local level. Only rarely can a group of teachers or local school administrators make drastic changes in these time allocations. And, since these time allocations are much the same for most students, they account for only small differences in the learning of individual students within a classroom or school.

Quite in contrast to the concept of time *available* for learning is the variable of *time-on-task* (i.e., active learning time, time that students are engaged in learning). If two students are in the same classroom and one is actively engaged in learning for 90% of the classroom hour while the other is actively engaged for only 30% of that hour, there will be quantitative as well as qualitative differences in their learning during that hour.

One method of appraising time-on-task is to determine at various intervals whether or not a particular student is *overtly* engaged in the learning — paying attention, doing work assigned, or in some way responding in a relevant way to the instruction and the instructional material. A second method is to determine the extent to which the student is *covertly* engaged in the learning. This is done by various methods (stimulated recall, inter-

---

BENJAMIN S. BLOOM is distinguished service professor of education, University of Chicago. This article is adapted from the introduction to his new book, All Our Children Learning, *scheduled for publication by McGraw-Hill this spring. © 1980, Benjamin S. Bloom.*

BEST COPY AVAILABLE

whether the student is *thinking* in relevant ways about what is going on in the classroom or whether his thoughts are unrelated to the classroom teaching/learning processes. Most studies report an index of time-on-task as the proportion of the classroom hour the individual student was on task — overtly, covertly, or an average of the two.

Studies of this variable show that the percentage of engaged time (for individual students or groups of students) is highly related to subsequent measures of achievement and to subsequent indices of interests or attitudes toward the learning. In turn, time-on-task is largely determined by the quality of instruction and the extent to which the students have the cognitive prerequisites for each new learning task. To put it another way, students cannot actively engage in learning if the instruction is poor and/or they are unable to comprehend what is being taught and what they are to do.

For the purposes of this article, what is most important is the strong evidence that the amount of active engaged time in the classroom can be altered during a sequence of learning tasks. Consider two groups of students who are comparable in aptitude or previous achievement at the beginning of a new course. One group learns the subject under conventional conditions while the second learns under a very high quality of instruction (mastery learning or some other procedure that maximizes learning). During the first learning task both groups are likely to be very similar in percentage of time-on-task. On the second learning task, the percentage of time-on-task will tend to be greater for the group with high-quality instruction and lower for the group with poorer instruction. If both groups are followed over a series of learning tasks, the high-quality instruction group will be found to increase greatly in percentage of time-on-task while the low-quality instruction group decreases greatly in percentage of time-on-task. On the final learning task the two groups (who were very similar on the first learning task) will be very different. These differences will be reflected in achievement differentials, motivation for further learning of the subject, and self-confidence in learning ability.

Time-on-task is then one of the variables that account for learning differences between students, between classes, and even between nations. Time-on-task can be altered positively (or negatively) by the instructional process, and this has direct consequences for the learning that will take place.

Intelligence vs. Cognitive Entry

During much of this century educators have used intelligence and aptitude tests to predict school achievement. In general, [illegible] later achievement have been found to be about +.30 to +.70. Most researchers and educators have interpreted these relations as indications that intelligence and aptitude *determine* the individual's potential for learning. Many educators use these test scores as a basis for making long-term decisions about selection, grouping, and even about the types of school programs to which individual students are assigned. All too frequently, intelligence and aptitude scores determine opportunities for further education, student support and encouragement, and even the types of interaction between teachers and students.

There is some evidence that intelligence test scores are alterable in the early years (ages 3-7), but there is little evidence of significant alteration in levels of intelligence as a result of school experiences in the later years. Less is known about the alterability of performance on specific aptitude tests. On the basis of present evidence, we may regard both intelligence and aptitude as highly stable characteristics.

Quite in contrast to intelligence and aptitude indices are *cognitive entry characteristics*. These are the specific knowledge, abilities, or skills that are essential prerequisites for the learning of a particular school subject or a particular learning task. Such prerequisites typically correlate +.70 or higher with measures of achievement in a subject. Furthermore, when they are identified and measured, they *replace* intelligence and aptitude tests in the prediction of later achievement. That is, intelligence or aptitude tests add little or nothing to cognitive entry measures for the prediction of school learning. All of this is to say that cognitive entry characteristics have a high relation to achievement, and they have an obvious causal effect on later achievement. This is especially true when sequential learning tasks are involved, where it may be impossible to learn task B without prior adequate learning of task A.

Cognitive entry characteristics are highly alterable, because they represent particular content and skills that may be learned if they are absent, reviewed if they have been forgotten, and learned to a criterion level if they have been learned to a lesser level. In the next section of this article I shall refer to feedback-corrective procedures as one major method for insuring that cognitive entry characteristics are developed adequately for nearly all students. Much of mastery learning research demonstrates that the large gains in final achievement for mastery versus control groups are attributable to the fact that the mastery students were brought to high levels of achievement on the *prerequisites* for each new learning task. This was not done for the control students.

Much of the variation in school learning is directly determined by the variation [illegible] cognitive entry characteristics. When means are found for insuring that students reach adequate levels of competence on the essential cognitive entry behaviors, most students can be assured of high levels of school learning with very little variation in their achievement. The alterability of cognitive entry characteristics has the most profound implications for instruction, curriculum, and our views about the learning potential of almost all students in the schools.

Summative vs. Formative Testing

In most classrooms achievement tests are used for summative purposes. The summative test evidence is primarily used to classify or judge the student on the extent to which he has learned the content and objectives set for the course. The students' scores on each test are converted into school marks or other indices that compare each student with norms or standards set by the teacher or the test makers. Typically, once a student has taken a test, he is marked and rarely is given an opportunity for correcting his errors or being retested. The basic notion is that the students have had equal opportunity to learn the subject over a defined period of time and are then judged on what they have learned. This is repeated again and again during the school year.

Test results and school marks are frequently assumed to be the primary motivators for learning in the school. Marks based on tests are also assumed to be sound estimates of the *quality of learning* as well as a proper index of the *quality of the learners*. Such marks are eventually the basis for many decisions about learners, including school programs and further opportunities for education.

The use of summative testing/grading procedures results in highly predictable measures of school achievement. Typically, correlations between achievement tests in the same subject at two points in time are above +.70 (depending upon the reliability of the separate tests). If carefully made standardized tests are used over a number of subjects, the correlations over a five-year period or longer tend to be +.80 or higher. That is, the ranks of the students in a school remain very constant over many years of schooling. Many researchers and educators infer from this that differentials in achievement are non-alterable and that they are fixed by intelligence, heredity, home influences, or other conditions outside the school. It is assumed that the student and his background explain this remarkable stability of achievement and that the causes or remedies are not to be found within the schools. It is the student who has failed (or succeeded), and the teacher, the instruction, the curriculum, or the school are not to be held responsible.

In contrast to tests used for grading

BEST COPY AVAILABLE

> "When formative tests and corrective procedures are used . . . as high as 80% or 90% of the students are able to reach the mastery standard. . . ."

and judging is the use of tests and other evidence as an integral part of the formation of the learning. Formative tests are used primarily as feedback, to inform the student about what he has learned well and what he still needs to learn. When feedback is provided in relation to corrective procedures to help the student correct the learning, then with additional time and help most students do reach the standard of achievement set by the teacher. Typically, teachers use a parallel formative test to determine when the student has completed the corrective process to the set standard. Various studies have found that if 20% of a group of students reach the mastery standard on a formative test given at the end of a particular learning task, then with an hour or two of corrective effort most of the students reach the same mastery standard, when they are retested on a parallel formative test.

When formative tests and corrective procedures are used in this way over a series of learning tasks, the proportion of students reaching the mastery standard (before correctives) increases on each subsequent task until as high as 80% or 90% of the students are able to reach the mastery standard on the final learning tasks in the series. The amount of corrective help needed becomes smaller on successive learning tasks, until only a few students need such corrective procedures. The students appear to be "learning to learn."

This use of formative tests insures that most of the students have the necessary cognitive prerequisites for each new learning task, that students have increased interest in learning and greater confidence in their own ability to learn, and that they use more of the classroom time to engage actively in the learning process.

Formative tests are also useful in helping the teacher determine which aspects of the learning task were learned well by the majority of the students and which were learned poorly by most of the students. This gives the teacher feedback in order to determine which ideas and skills need to be reviewed or retaught in a different way if the majority of students are to learn them to a high standard. The major change is that teachers do less in the way of judging and grading students on what they had learned by a particular date and they do more to see to it that each student learns what he or she needs as preparation for the next learning task(s).

Formative testing in relation to the corrective process may be considered as an example of cybernetic feedback-corrective procedures necessary for almost all human activities. In tutoring situations the one-to-one relation provides so many cues that the feedback-corrective process is almost a natural interactive exchange between tutor and tutee. However, group learning is central in the schools, and it is difficult to provide adequate feedback-correctives for the teacher and the 30 or so learners in each classroom. As a result, much teaching may take place with inadequate learning on the part of many of the students. Periodic formative testing and corrective procedures can be effective as one way of insuring that excellent learning takes place. However, in the long run, the basic problem of group learning is to find ways of providing feedback-corrective processes as an integral part of the classroom teaching/learning interactions.

Teachers vs. Teaching

Over the past four decades there has been a great deal of research on teacher characteristics and their relations with student learning. This research has been concerned with such variables as the age of the teachers, their training, teaching experience, membership in teacher organizations, personality and attitudes, and even performance on achievement tests related to their field of teaching. In general, the relationship between teacher characteristics and student learning has typically been represented by correlations of less than +.20. Researchers in the past may not have selected the right teacher characteristics for study. However, based on the research done to date, we may conclude that the characteristics of teachers have little to do with the learning of their students. And, even if they did show higher relations, most of the characteristics of teachers studied so far are static variables not directly alterable by inservice or other teacher training programs.

Different from these many studies of teacher characteristics is the more recent research on the qualities of teaching that have a direct causal relation with student learning in the classroom. This research on the qualities of *teaching* (rather than on qualities of the *teachers*) consists largely of observational and experimental studies of teachers interacting with their students. Although there are many ways of doing this research, the theoretical approach of John Dollard and Neal Miller has been found very useful. Dollard and Miller have emphasized three major characteristics of all teaching: *cues, reinforcement*, and *participation.*[*] *Cues* include instruction as to what is to be learned as well as directions as to what the learner is to do in the learning process. Much of the research relates student learning to the clarity, variety, meaningfulness, and strength of the explanations and directions provided by the teacher and/or the instructional material. *Reinforcement* includes the extent to which the student is rewarded or reinforced in his learning. Much of the research relates student learning to the variety of reinforcements provided, the frequency with which reinforcement is used, and the amount and kind of reinforcement given to different students in the class. *Participation* includes the extent to which the student actively participates or engages in the learning. The research relates student learning to the extent to which he actively participates in using the cues, makes appropriate responses, and practices the responses until they have become a part of his repertoire. The research also includes the extent to which the instructor and/or the instructional method engages the different students in the class in overt as well as covert participation and response to the learning.

Observations of teacher interaction with students in the classroom reveal that teachers frequently direct their teaching and explanations to some students and ignore others. They give much positive reinforcement and encouragement to some students but not to others, and they encourage active participation in the classroom interaction from some students and discourage it from others. The studies find that typically the students in the top third of the class are given the greatest attention by teachers, while the students in the bottom third receive the least attention and support. These differences in the interaction between teachers and students provide some students with much greater opportunity and encouragement for learning than is provided other students in the same classroom.

These qualities of teaching are alterable as a result of inservice education that provides teachers with feedback on what they are doing (or not doing) and what they can do to alter the situation. Studies have found that when these interactions of teachers with their students are altered, there are significant improvements in student learning.

---

[*] In *Personality and Psychotherapy* (New York: McGraw-Hill, 1950).

BEST COPY AVAILABLE

[illegible] favorable conditions of learning for some students in the class than for others. Generally, they are under the impression that all students are given equal opportunity for learning. When teachers are helped to secure a more accurate picture of their own teaching methods and styles of interaction with their students, they are increasingly able to provide favorable learning conditions for most of their students, rather than just for the top fraction of the class.

As the qualities of teaching become more central than the characteristics of the teachers, we shall see more clearly the kinds of training that can improve teaching and learning. We shall also see more clearly the variety of conditions that can serve in the teaching/learning process. Tutors, aides, parents, and even other students can be helpful in promoting learning. New materials and media, new organization of classrooms, and new relations between teachers and students will be seen as important in contributing to learning.

Parent Status vs. Home Environment

Teachers and researchers have long known that children from some homes learn better in school than children from other homes in the same communities. In general, learning in the schools has been found to be related to the education and occupations of the parents, to the social class and socioeconomic status of the parents, and to membership in particular ethnic groups and races. Studies of socioeconomic status (which include parent education, occupation, and income) reveal correlations of +.30 to +.50 between such indices and measures of school achievement. While such studies do demonstrate significant effects of the home on school achievement, they are not very helpful to the schools or the parents, because these characteristics are not alterable. There is little the school or the parents can do to alter their level of education, occupation, income, or ethnic characteristics. While such studies may be of some slight value for predicting levels of learning of groups of children, they offer no specific clues as to what the schools or the parents can do to improve their children's learning.

In contrast to these earlier studies of the characteristics of parents are the more recent studies that emphasize what parents *do* in interacting with their children. These studies use interviews and observational techniques to investigate the environmental process variables in the homes — the interactions between children and their parents. Some of the home environmental process variables that appear to be most significant are: contribution of the home to development of the mother tongue, encouragement of the parents for [illegible] children, provision of help in learning when the child most needs it, and ways in which time and space are organized in the home. Such variables, when combined, correlate +.70 to +.80 with measures of school achievement. In general, the correlations are highest with school achievement involving reading, vocabulary, and problem solving and lowest with spelling and arithmetic computation. These results suggest that the home has greatest influence on the language development of the child, his general ability to learn, and his motivation to learn well in school. The home has least influence on specific skills taught primarily in the school.

It is clear that when the home and the school have congruent learning emphases, the child has little difficulty in his later school learning. But when the home and the school have divergent approaches to life and to learning, the child is likely to be penalized severely by the school — especially when school attendance is required for 10 or more years.

In recent years a large number of studies have attempted to alter some of these process variables in the home. These studies have made use of home visitors, special courses for parents, and parent involvement in the schools for brief periods [illegible] visual and written materials and games to be used at various points in the child's development. The research makes it clear that many of these process variables are alterable, and the effects of such alterations on the children's school learning are very great.

Even when these variables cannot be altered in the home, a knowledge of the home environmental process variables furnishes a basis for programs of early childhood and primary education that can compensate for particular deficiencies in the home. The major point to be made is that there is a curriculum and teaching style in each home and that it is the variations in the home curriculum and teaching that account for much of the difference in children's preparation for the learning tasks in the school.

"Yes, my child has learned a lot from public TV! She's learned fluent Spanish and perfect metrics, neither of which her mother understands."

Summary

If we are convinced that little or nothing can be done to improve the learning of individual students, then our major effort must be invested in predicting school achievement and classifying children at an early age. Stable variables are ideal for this purpose. Such efforts result in a school system that is quite effective for a small proportion of the students while at the same time it dooms most students to a deep sense of inadequacy and a dislike for school and school learning. Such a school system must invest much in the way of human and material resources with very small returns to the society or to the majority of its students.

If we are convinced that a good education is necessary for all who live in modern society, then we must search for the alterable variables that can make a difference in the learning of children and adults in or out of the school. Such alterable variables will do much to explain the learning process, and they will do even more to directly improve the teaching and learning processes in the schools. Our basic research task is to further understand how such alterable variables can be altered and their consequent effect on students, teachers, and learning.

The small number of alterable variables I have discussed here are only a few of the variables that have already been studied by researchers and used by teachers. These have already made a great difference in our understanding of school learning. But, also, they have brought about major changes in our views of learners and their amazing potential for learning. I hope that this small list will be rapidly expanded in the next decade and that they will become equally central for teachers, parents, and researchers. When they are thoroughly understood and well used, they will bring about the most profound changes in the schools and in the society. ☐

BEST COPY AVAILABLE